

Non-Gaussianity and extragalactic foregrounds to the Cosmic Microwave Background

Université Paris-Sud

THÈSE DE DOCTORAT

soutenue le 23/9/2013

par

Fabien Lacasa

Directrice de thèse : Nabila Aghanim Institut d'Astrophysique Spatiale, Orsay

Composition du Jury

<i>Président :</i>	Jean-Loup Puget	Institut d'Astrophysique Spatiale, Orsay
<i>Rapporteurs :</i>	Olivier Doré	Jet Propulsion Laboratory
	Jean-Philippe Uzan	Institut d'Astrophysique de Paris
<i>Examineurs :</i>	Francis Bernardeau	Commissariat à l'énergie atomique
	Bartjan van Tent	Laboratoire de Physique Théorique, Orsay
<i>Membres invités :</i>	Martin Kunz	Université de Genève
	Guilaine Lagache	Institut d'Astrophysique Spatiale, Orsay

Abstract

This PhD thesis, written in english, is interested in the non-Gaussianity of extragalactic foregrounds to the Cosmic Microwave Background (CMB). Indeed, the CMB is one of the golden observables of contemporary cosmology, allowing in particular to highly constrain the cosmological parameters describing the universe. In the last decade has emerged the interest of looking for deviations of the CMB statistics to the Gaussian law. Indeed, the standard models for the generation of primordial perturbations (inflation) predict a close to Gaussian statistic, where the deviations to this law would in fact allow to discriminate the different models. However the measurements of the CMB, in particular by the *Planck* satellite for which I worked, can be contaminated by several emissions also known as foregrounds, depending on the observing frequency. I got interested in particular in some of these foregrounds due to extragalactic emissions tracing the large scale structure of the universe, specifically to radio and infrared point-sources (the latter constituting the Cosmic Infrared Background, CIB) and to the thermal Sunyaev-Zel'dovich effect (tSZ).

In this thesis, I hence describe the tools to characterise a random field and its deviation to the Gaussian law. The harmonic space is well adapted for theoretical and numerical computations and I thus define polyspectra. This thesis is particularly interested in the third order polyspectrum, the bispectrum, as it is the lowest order indicator of non-Gaussianity, with potentially the highest signal to noise ratio. I describe how the bispectrum can be estimated on data, accounting for the complexity due to partial sky coverage (because of a galactic mask for example) and instrumental resolution and noise. I then propose a method to visualise the bispectrum, which is more adapted than the already existing ones as it allows to probe the scale and configuration dependence while accounting for the invariance under the permutation of multipoles. I then describe the error bars/covariance of a polyspectrum measurement, a method to generate non-Gaussian simulations (useful to test the effect of diverse instrumental complications on a measurement), and how the statistic of a 3D field projects onto the sphere when the observations are line-of-sight integrals.

I then motivate my research from the cosmological point of view, by describing the generation of density perturbations ¹ by the standard inflation model and their possible non-Gaussianity, how they yield the Cosmic Microwave Background anisotropies and grow to form the large scale structure of today's universe. To describe this large scale structure, I present the halo model and propose a diagrammatic method allowing to compute the polyspectra of the galaxy density field and to have a simple and powerful representation of the involved terms.

I then describe the different foregrounds to the Cosmic Microwave Background, galactic as well as extragalactic. I briefly describe the physics of the thermal Sunyaev-Zel'dovich effect and explain how its spatial distribution can be described with the halo model. I then describe the extragalactic point-sources and present a prescription for the

¹called adiabatics, the possibility of isocurvature modes being neglected here.

non-Gaussianity of clustered sources and its generalisation to the case of several populations of sources, clustered or not. For the CIB I introduce a physical modeling with the halo model, and the diagrammatic method proves to be particularly useful in this case. I implement numerically the model to compute the 3D galaxy bispectrum and to produce the first theoretical prediction of the CIB angular bispectrum. I then explore the model results : I show the contributions of the different terms and the temporal evolution of the galaxy bispectrum. For the CIB angular bispectrum, I show its different terms and its scale and configuration dependence. I also show how the bispectrum varies with model parameters and which constraints a measurement would bring to these parameters. In the considered case, the bispectrum allows very good constraints, as it either breaks degeneracies present at the power spectrum level or either these constraints are better than those coming from the power spectrum.

Finally, I describe my work on measuring non-Gaussianity. I first explain the estimator that I introduced for the amplitude of the CIB bispectrum, and how this estimator can be combined with similar ones for radio sources and the CMB to produce a joint and robust constraint of the different sources of non-Gaussianity. Then, I quantify the contamination that extragalactic point-sources can produce to the estimation of primordial NG ; in the *Planck* case I show that this contamination is negligible for the frequencies where the CMB is dominant. I then describe my measurement of the CIB bispectrum on *Planck* data : the different difficulties encountered and the obtained results. The bispectrum is very significantly detected at 217, 353 and 545 GHz with signal to noise ratios ranging from 5.8 to 28.7. Its shape is consistent between frequencies, as well as the intrinsic amplitude of NG, which guarantees the results' robustness. Interestingly, the measured bispectrum is significantly steeper than the empirical prescription that I introduced, and than the physical model based on the halo model described precendently. I speculate that this difference indicates the necessity to complexify the infrared emission model, and that it can be explained by the most recent models which consider that the emissivity depends on the host halo mass and on the type of galaxy (central/satellite). Ultimately, I describe my measurement work on the thermal Sunayev-Zel'dovich bispectrum, on simulations and on Compton parameter maps estimated by *Planck*. For the *Planck* data, the robustness of the estimation is validated thanks to realist foreground simulations, which allow to quantify their contamination to the estimated tSZ map. The tSZ bispectrum is very significantly detected with a signal to noise ratio ~ 200 , and I show that its amplitude is consistent with the projected map of detected clusters and with the *Planck* Sky Model simulation ; I also show that its scale and configuration dependence is consistent with that I found on tSZ simulations. Finally, I use this measurement to put a constraint on the cosmological parameters σ_8 and Ω_b , obtaining $\sigma_8 (\Omega_b/0.049)^{0.35} = 0.74 \pm 0.04$ in agreement with the estimations of σ_8 through other tSZ statistics.

This thesis lead to four articles : Lacasa et al. (2012) published by MNRAS, Lacasa & Aghanim (2013) accepted by Astronomy & Astrophysics and being revised, and Lacasa et al. (2013) & Pénin et al. (2013) accepted by MNRAS.

I also contributed to the following *Planck* articles : Planck Collaboration XXI (2013), Planck Collaboration XXIV (2013) and Planck Collaboration XXX. (2013).

Contents

Introduction	1
1 Some statistics	3
1.1 Random variables	4
1.1.1 Notations	4
1.1.2 Moments	4
1.1.3 Cumulants	5
1.1.4 Examples and limitations	6
1.1.4.1 Gaussian law	6
1.1.4.2 Poisson law	7
1.1.4.3 Probability laws without moments nor cumulants	7
1.1.4.4 Probability law not uniquely defined by its moments or cumulants	8
1.2 Random fields	9
1.2.1 Definition	9
1.2.2 Correlation functions	9
1.2.3 Ursell functions – or connected correlation functions	10
1.2.4 Gaussian random fields	11
1.2.5 White noise	11
1.2.6 Gram-Charlier expansion	12
2 Some spherical analysis	14
2.1 Spherical harmonics	16
2.1.1 Definition	16
2.1.2 Useful properties	16
2.1.3 Relation with angular momentum and Gaunt coefficients	17
2.2 Harmonic coefficients and polyspectra	18
2.2.1 Harmonic coefficients	18
2.2.2 Isotropic polyspectra	18
2.2.3 Relation with real-space correlation functions	20
2.2.4 Estimation of the power spectrum and bispectrum	21
2.2.4.1 Power spectrum	21
2.2.4.2 Bispectrum	23
2.2.4.3 Effect of masking on the bispectrum : variance	25
2.2.4.4 Effect of masking on the bispectrum : bias	27
2.2.4.5 Binned bispectrum coupling matrix evaluation	28
2.2.4.6 Bispectrum representation : a new parametrisation	32

2.2.5	(Co)variance	32
2.2.5.1	Power spectrum	33
2.2.5.2	Bispectrum	34
2.2.5.3	Incomplete sky	37
2.3	Examples	37
2.3.1	Gaussian random field	37
2.3.2	White noise	38
2.3.2.1	Generating non-Gaussian fields	39
2.4	Relation with Fourier statistics	39
2.4.1	2D polyspectra	40
2.4.2	3D polyspectra	41
2.4.3	Plane wave expansion	42
2.4.4	Projection of the past light cone	42
3	Some cosmology	45
3.1	Friedman-Lemaître space-time	46
3.2	Inflation	47
3.2.1	General paradigm	47
3.2.2	Single field inflation	48
3.2.2.1	Background evolution	48
3.2.2.2	Perturbation generation and primordial non-Gaussianity	49
3.3	Cosmic Microwave Background anisotropies	50
3.4	Large Scale Structure	54
3.4.1	Linear growth	54
3.4.2	Local bias scheme	56
3.4.3	Halo model	58
3.4.4	A diagrammatic method for galaxy correlation functions	61
4	Extragalactic foregrounds : tracers of the Large Scale Structure	66
4.1	Foregrounds : introduction	67
4.2	Sunyaev-Zel'dovich effect	71
4.2.1	Introduction : ionisation and photon scattering	71
4.2.2	Description with the halo model	72
4.3	Extragalactic point-sources	74
4.3.1	Introduction	74
4.3.2	Prescription for the non-Gaussianity of clustered and unclustered sources	78
4.3.3	CIB description with the halo model	81
5	Measuring non-Gaussianity	93
5.1	Observational considerations for <i>Planck</i>	94
5.2	Non-Gaussianity from extragalactic point-sources	95
5.2.1	IR amplitude and joint NG constraints	95
5.2.2	Contamination of primordial NG estimation	101
5.2.3	Measuring the CIB NG with <i>Planck</i> data	105
5.2.3.1	Small fields	106
5.2.3.2	GASS field : bispectrum estimation pipeline	107

5.2.3.3	GASS field : results and post-treatment	111
5.2.3.4	GASS field : exploitation of the measurements	115
5.3	Measuring the thermal Sunyaev-Zel'dovich NG	117
5.3.1	On simulations	117
5.3.2	<i>Planck</i> tSZ data : characteristics and processing pipeline	118
5.3.3	<i>Planck</i> tSZ data : results and characterisation	123
Conclusion		127
Perspectives		129
Glossary		131

Introduction

Sky observations are nowadays carried from radio waves to gamma rays, extending astronomy across the whole electromagnetic spectrum. In particular, space-based all-sky surveys in the microwave range have become available since RELIKT-1 (Klypin et al. 1987), and later on with COBE (Boggess et al. 1992), WMAP (Bennett et al. 1997), and now *Planck* (Planck Collaboration et al. 2011d). These missions were designed to map the Cosmic Microwave Background (CMB), a relic radiation from the early universe emitted when it was in a plasma state. The Cosmic Microwave Background is indeed one of the golden observables of contemporary cosmology, along with standard rulers such as type Ia supernovae (Riess et al. 2009), and Large scale Structure (LSS) tracers, e.g. allowing the measurement of Baryon Acoustic Oscillations (Percival et al. 2010).

Extraction of cosmological information from the observed CMB map involves the measurement of its spatial correlation functions, in particular at second order i.e. the power spectrum. For example, the *Planck* mission has put the stringest constraint to date on cosmological parameters using the measurement of this power spectrum (Planck Collaboration XVI 2013). However, in the last decade interest has risen for the measurement of higher orders (non-Gaussianity studies) as it may provide information on the primordial process generating the cosmological perturbation, in particular it may discriminate inflation models which are degenerate at the power spectrum level (Bartolo et al. 2004).

Foreground signals are nevertheless present at microwave frequencies and contaminate the CMB observations. These foregrounds stem from the emission of LSS tracers and probe information of cosmological interest (e.g., the ionised gas in clusters or the star-formation history). High order measurements are also of interest for these LSS tracers. Indeed, they allow for a more complete statistical characterisation of the field, hence an extraction of more cosmological information.

In the literature, the study of foregrounds non-Gaussianity first focussed on the case of extragalactic radio point-sources, which are the dominant non-Gaussian signal outside the galactic plane at WMAP frequencies. These studies were used to quantify the contamination of point-sources to the CMB power spectrum (Komatsu et al. 2003) or to the estimation of primordial non-Gaussianity (e.g. Argüeso et al. 2003; González-Nuevo et al. 2005; Babich & Pierpaoli 2008). The study of LSS non-Gaussianity has also emerged in the case of galaxy surveys. It has indeed been shown to be a powerful probe cosmological parameters as well as galaxy formations model, in a way complementary to the power spectrum (see e.g. the thesis by Sefusatti 2005).

This thesis is interested in the study of high order correlation functions for different extragalactic foregrounds to the CMB, in particular at third order. This study has two main motivations. First, it is required for an unbiased measurement of the corresponding

information for the primordial CMB. Second, the measurement of these high orders and its comparison with theoretical expectation allows a more complete statistical analysis of the foreground signals. This statistical analysis could in turn provide better constraints on the models of these foregrounds, and thus a better understanding of the processes they trace and of the LSS distribution.

The thesis is divided as follows : in the first chapter, I introduce the statistical tools needed to study random fields.

The second chapter focuses on the case of random field on the sphere, such as sky observations, and introduces harmonic analysis. It discusses in particular the issue of the estimation of high orders, and the projection of these statistics from the 3-dimensional space, where models are naturally set, to the celestial sphere.

The third chapter introduces the basics of cosmology motivating this thesis, and shows how the primeval perturbations of the universe are generated during the inflation period. It further shows how these perturbations generate the CMB anisotropies and later on the Large Scale Structure distribution. In particular I describe a diagrammatic formalism I have developed, which allows the computation of the high order correlation functions of the galaxy density field, a standard tracer of the LSS.

The fourth chapter introduces the different extragalactic foregrounds to the CMB that trace the LSS, in particular the Cosmic Infrared Background (CIB) and the thermal Sunyaev-Zel'dovich (tSZ) signal. Furthermore it shows how their correlation functions can be modeled theoretically.

The last chapter describes the data analysis I have performed for *Planck* on these extragalactic foregrounds. It shows my estimations of how CMB high order studies are biased by extragalactic foregrounds, and how to account for this bias. It also shows the third order measurements that I have produced for the CIB and tSZ signals, and the constraints that I could derive from these measurements.

A glossary is given after the conclusion and the perspectives, listing abbreviations used and symbols that may not be standard.

This thesis lead to four articles : Lacasa et al. (2012) published by MNRAS, Lacasa & Aghanim (2013) accepted by Astronomy & Astrophysics and being revised, and Lacasa et al. (2013) & Pénin et al. (2013) submitted to MNRAS.

I furthermore contributed to the following *Planck* articles : Planck Collaboration XXI (2013), Planck Collaboration XXIV (2013) and Planck Collaboration XXX. (2013).

Chapter 1

Some statistics

Contents

1.1	Random variables	4
1.1.1	Notations	4
1.1.2	Moments	4
1.1.3	Cumulants	5
1.1.4	Examples and limitations	6
1.1.4.1	Gaussian law	6
1.1.4.2	Poisson law	7
1.1.4.3	Probability laws without moments nor cumulants	7
1.1.4.4	Probability law not uniquely defined by its moments or cumulants	8
1.2	Random fields	9
1.2.1	Definition	9
1.2.2	Correlation functions	9
1.2.3	Ursell functions – or connected correlation functions	10
1.2.4	Gaussian random fields	11
1.2.5	White noise	11
1.2.6	Gram-Charlier expansion	12

In this chapter, which has no pedagogical pretention, I describe the statistical characterisation of random objects. I begin with the simpler case of random variables, before generalising to random fields. The statistical notions described are illustrated on several examples of distributions of random variables or random fields, and these examples find applications later on in the thesis.

1.1 Random variables

1.1.1 Notations

Let X be a random element, valued on some space E . Its probability law is a positive measure on E , with total measure unity :

$$\int dP_X = 1 \quad (1.1)$$

The probability of an event A is :

$$\mathbb{P}(X \in A) = dP_X (X^{-1}(A)) \quad (1.2)$$

For circumstances of interest in cosmology, the probability law derives from a *probability density function* (p.d.f.) :

$$dP_X = p(x) dx \quad (1.3)$$

with $p(x)$ a positive function (with integral unity) and dx the Lebesgues measure on the space of interest, that is \mathbb{R}^n or \mathbb{C}^n throughout this thesis.

Any function $g(X)$ then has an expectation value¹ :

$$\langle g(X) \rangle = \int g(x) p(x) dx \quad (1.4)$$

if the integral converges.

Two random elements are said *independent* (noted $X \perp\!\!\!\perp Y$) iff² for any functions g and h :

$$\langle g(X) h(Y) \rangle = \langle g(X) \rangle \langle h(Y) \rangle \quad (1.5)$$

A random variable is a scalar random element, i.e. it is valued in \mathbb{R} or \mathbb{C} .

1.1.2 Moments

The n -th order moment of a random variable X is defined, for $n \in \mathbb{N}$, as :

$$m_n = \langle X^n \rangle \quad (1.6)$$

if the involved integral converges.

In particular m_1 is called the mean (or average) of X . One may define the *centered moments* by substracting this average :

$$\mu_n = \langle \tilde{X}^n \rangle \quad \text{with} \quad \tilde{X} = X - \langle X \rangle \quad (1.7)$$

¹traditionally noted $\mathbb{E}[g(X)]$ in mathematics, but we use here lighter notations in agreement with cosmology's conventions.

²if and only if

In the following, we will always consider centered moments and loosely refer to them as just ‘moments’ for concision.

We note in passing that moments may not be defined beyond a certain order, explicit examples being given in Sect.1.1.4.

The *moment-generating function* (m.g.f.) is defined, for $t \in \mathbb{R}$, as :

$$M_X(t) = \langle e^{tX} \rangle \quad (1.8)$$

if the expectation value exists. In other terms, it is the Laplace transform of the p.d.f. It may be defined only at $t = 0$ for some random variables (see Sect.1.1.4). However, if it can be defined on \mathbb{R} and is sufficiently regular, the p.d.f. can be recovered from the moment-generating function through an inverse Laplace transform, meaning that two random variables having the same moment-generating function have the same probability law.

The moment-generating function inherits its name from the following property : if it is defined at least on a neighbourhood of $t = 0$, then :

$$m_n = \left. \frac{d^n M_X(t)}{dt^n} \right|_{t=0} \quad (1.9)$$

and, if the series converges,

$$M_X(t) = \sum_{n=1}^{+\infty} \frac{m_n t^n}{n!} \quad (1.10)$$

This is an important result for statistics as it means that the study of a random variable can be reduced to the measurement of the hierarchy of its moments (if the series in Eq.1.10 converges). So that the uncountably infinite problem of measuring the p.d.f. of a random variable is reduced to the countable (but still infinite) problem of measuring its moments.

In practice one may want to measure only a finite number of moments and truncate the series Eq.1.10 at that order. This approach will be further developed and discussed in Sect.1.2.6.

The moment-generating function has the following property :

$$X \perp\!\!\!\perp Y \Leftrightarrow M_{X+Y} = M_X M_Y \quad (1.11)$$

which will be useful later, in this chapter.

1.1.3 Cumulants

For a random variable, the *cumulant-generating function* (c.g.f.) is defined by :

$$g_X(t) = \log(M_X(t)) \quad (1.12)$$

which defines the cumulants as :

$$\kappa_n = \left. \frac{d^n g_X(t)}{dt^n} \right|_{t=0} \quad (1.13)$$

so that :

$$g_X(t) = \sum_{n=1}^{+\infty} \frac{\kappa_n t^n}{n!} \quad (1.14)$$

Under the previously mentioned regularity hypothesis, the cumulant-generating function characterises univoquely the probability law. Hence characterising a random variable can be achieved by studying its cumulants.

Through Eq.1.11 we have $g_{X+Y} = g_X + g_Y$ iff $X \perp\!\!\!\perp Y$, hence cumulants have the following property :

$$X \perp\!\!\!\perp Y \quad \Leftrightarrow \quad \kappa_n(X+Y) = \kappa_n(X) + \kappa_n(Y) \quad \forall n \quad (1.15)$$

This property is not shared by (centered) moments, hence cumulants are more advantageous for theoretical or practical computations.

Cumulants are related univoquely and hierarchically to moments : the cumulant at a given order is a function of the moments up to that order, and vice-versa. The following equations give these relations up to order 6, as they will be useful later on :

$$\kappa_1 = m_1 \quad (1.16)$$

$$\kappa_2 = \mu_2 \quad (1.17)$$

$$\kappa_3 = \mu_3 \quad (1.18)$$

$$\kappa_4 = \mu_4 - 3\mu_2^2 \quad (1.19)$$

$$\kappa_5 = \mu_5 - 10\mu_3\mu_2 \quad (1.20)$$

$$\kappa_6 = \mu_6 - 15\mu_4\mu_2 - 10\mu_3^2 + 30\mu_2^3 \quad (1.21)$$

and conversely :

$$m_1 = \kappa_1 \quad (1.22)$$

$$\mu_2 = \kappa_2 \quad (1.23)$$

$$\mu_3 = \kappa_3 \quad (1.24)$$

$$\mu_4 = \kappa_4 + 3\kappa_2^2 \quad (1.25)$$

$$\mu_5 = \kappa_5 + 10\kappa_3\kappa_2 \quad (1.26)$$

$$\mu_6 = \kappa_6 + 15\kappa_4\kappa_2 + 10\kappa_3^2 + 15\kappa_2^3 \quad (1.27)$$

Terminology : κ_2 is called the variance, often noted σ^2 , and its square root σ is called the standard deviation. It is sometimes used to normalise higher order cumulants, e.g. the skewness is defined as κ_3/σ^3 , and the (excess) kurtosis is κ_4/σ^4 . In this thesis, we will not use this normalisation unless explicitly mentioned, so we may loosely use the term skewness to refer to the third-order cumulant.

1.1.4 Examples and limitations

1.1.4.1 Gaussian law

The Gaussian (or normal) law is a real distribution (i.e. $X \in \mathbb{R}$) which is ubiquitous in physics, particularly because of the Central Limit theorem³. It is characterised by two

³which essentially tells that the average of many independent random variables tends to be Gaussian, for example the power deposited by photons on a *Planck* detector, as there are many photons hitting the detector within its response time.

parameters which can be chosen as its mean and variance, it is noted $\mathcal{N}(\mu, \sigma^2)$, and has p.d.f. :

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad \forall x \in \mathbb{R} \quad (1.28)$$

As the Laplace transform of a Gaussian is a Gaussian, the moment-generating function is calculated easily :

$$M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right) \quad (1.29)$$

Consequently the cumulant-generating function is :

$$g_X(t) = \mu t + \frac{\sigma^2 t^2}{2} \quad (1.30)$$

and the cumulants are :

$$\kappa_1 = \mu \quad (1.31)$$

$$\kappa_2 = \sigma^2 \quad (1.32)$$

$$\kappa_n = 0 \quad \forall n \geq 3 \quad (1.33)$$

The moments are more complicated and may e.g. be recovered through the formulas 1.22-1.27. The simplicity of cumulants is one of their conceptual and practical advantage and, anticipating on the following, it shows that studying cumulants at order ≥ 3 probes potential non-Gaussianity.

1.1.4.2 Poisson law

The Poisson law is a discrete distribution, meaning that the outcome of the random variable is an integer. It arises when counting the number of independent events happening on a given period of time, for example the number of photons hitting a *Planck* detector within one second. It is characterised by a single parameter λ being the mean, and the probability law is :

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad (1.34)$$

The m.g.f. is :

$$M_X(t) = \exp(\lambda(e^t - 1)) \quad (1.35)$$

so that the cumulants are :

$$\kappa_n = \lambda \quad \forall n \in \mathbb{N}^* \quad (1.36)$$

In particular the dimensionless skewness is $\kappa_3/\sigma^3 = \lambda^{-1/2}$ and can be made arbitrarily important by taking λ small enough. This property will be important in Sect.2.2.4.5 to generate non-Gaussian simulations.

1.1.4.3 Probability laws without moments nor cumulants

The Cauchy-Lorentz (Hazewinkel 2001) law is a real-valued distribution with a p.d.f. given by :

$$p(x) = \frac{1/\pi}{1+x^2} \quad (1.37)$$

Because the p.d.f. decreases slowly at infinity, all the moments, and hence cumulants, are undefined. Indeed the involved integral does not converge. The moment-generating function is also undefined for $t \neq 0$.

Furthermore, one may define a probability law such that the moments and cumulants are defined only up to a certain order n , e.g. :

$$p(x) \propto \frac{1}{1 + |x|^{n+2}} \quad (1.38)$$

defines a p.d.f. up to a normalisation factor. Its moment-generating function is also undefined for $t \neq 0$.

High order moments or cumulants are hence sensitive to the tail of the p.d.f., i.e. to rare events. A sufficient condition for the existence of cumulants is that the p.d.f. has bounded support or decreases faster than any rational fraction (e.g. as an exponential) at infinity.

1.1.4.4 Probability law not uniquely defined by its moments or cumulants

A random variable X follows a *log-normal* distribution if $Y = \log(X)$ follows a Gaussian distribution. It is noted $\ln \mathcal{N}(\mu, \sigma^2)$, with μ and σ^2 respectively the mean and variance of Y , and its p.d.f. is :

$$p(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \quad (1.39)$$

All the moments and cumulants are defined, namely the uncentered moments are :

$$m_n = \exp\left(\mu n + \frac{\sigma^2 n^2}{2}\right) \quad (1.40)$$

while the centered moments and the cumulants have more complicated expressions.

However the moment-generating function is undefined for $t \neq 0$ as the involved integral does not converge. Indeed the series in Eq.1.10 diverges, as $\frac{m_n}{n!} \rightarrow +\infty$ when $n \rightarrow +\infty$.

Hence we are not in the position of deducing the p.d.f. from the moments, and indeed it can be shown that there are other distributions sharing exactly the same moments (see Stoyanov (1987) Sect. 11.2 for explicit details). The log-normal distribution is called *moment-indeterminate*, in other terms statistical information is escaping the hierarchy of moments. This problem arises because the distribution is “fat-tailed”.

The inadequacy of the direct moment approach in this case has been proven of interest for cosmology by Carron (2011), as the matter density field approximately follows a log-normal distribution. However, taking a log transform of the field suffices to make the moment approach valid again, as the variable then becomes Gaussian distributed. This ‘Gaussianization’ process –i.e. a 1-point non-linear remapping such that the 1-point marginal becomes Gaussian or close to Gaussian– was first proposed for Large Scale Structure studies by Weinberg (1992) and further developed by Neyrinck (2011a,b). Note however that ‘Gaussianizing’ a general random field may not yield a Gaussian field⁴, so that the power spectrum of the Gaussianized field is not a complete

⁴For example, consider the p.d.f. with two points $p_\epsilon(x_1, x_2) \propto \text{Gauss}(x_1, x_2) (1 + \epsilon \sin(x_1 x_2))$ $\forall \epsilon \in [-1, 1]$, with ‘Gauss’ any two-point symmetric Gaussian. This p.d.f. is isotropic and its 1-point marginal is Gaussian, however it does not represent a Gaussian field.

statistical characterisation of the field. Finally, building on the same principle as the ‘Gaussianization’, several articles have been published during the writing of this thesis on applying non-linear transformations to put back information into the moment’s hierarchy of cosmological fields (Carron & Szapudi 2013; Leclercq et al. 2013; Simpson et al. 2013).

1.2 Random fields

1.2.1 Definition

A random field is a collection of random variables $\mathbf{X} = (X_i)_i$ with an index i running on a continuous space. Examples of relevance for cosmology are :

- the intensity of a signal depending on the direction in the sky. The index i is then running on the sphere and may be taken as the unit vector \hat{n} .
- the baryonic matter density (or velocity, pressure...) depending on the space-time point. The index i is then running on the manifold and may be taken as (\vec{x}, t) given a global choice of coordinates.

In the following, we will consider that the index runs over a finite number of values, as this will avoid technicalities such as dealing with functional derivatives. This is also realistic for practical applications as sky observations do not produce continuous fields but pixelised maps, with a finite number of pixels.

In summary, \mathbf{X} can be thought of as a large random vector, e.g. $X_i = T(\hat{n}_i)$ the temperature of a sky signal in direction \hat{n}_i with $i = 1 \cdots n_{\text{pix}}$.

The definition of independence in Sect.1.1.1 also applies to random fields.

1.2.2 Correlation functions

In the following sections, we will assume the index i of $\mathbf{X} = (X_i)_i$ runs from 1 to n . The correlation function (c.f. hereafter) of order k is then :

$$\mu_k(X_{i_1}, \cdots, X_{i_k}) = \left\langle \tilde{X}_{i_1} \cdots \tilde{X}_{i_k} \right\rangle \quad (1.41)$$

with $\tilde{X}_j = X_j - \langle X_j \rangle$, and with the usual hypothesis that the expectation value is well-defined.

The correlation functions are generated by the moment-generating function :

$$M_{\tilde{\mathbf{X}}}(\mathbf{t}) = \left\langle \exp \left(\mathbf{t} \cdot \tilde{\mathbf{X}} \right) \right\rangle \quad (1.42)$$

with $\mathbf{t} = (t_i)_i$ a vector, and the dot \cdot the canonical scalar product. Indeed, if the m.g.f. is defined, we have :

$$M_{\tilde{\mathbf{X}}}(\mathbf{t}) = \sum_{k=1}^{+\infty} \sum_{i_1 \cdots i_k=1}^n \frac{t_{i_1} \cdots t_{i_k}}{k!} \times \mu_k(X_{i_1}, \cdots, X_{i_k}) \quad (1.43)$$

and :

$$\mu_k(X_{i_1}, \dots, X_{i_k}) = \left. \frac{\partial^k M_{\mathbf{X}}(\mathbf{t})}{\partial t_{i_1} \dots \partial t_{i_k}} \right|_{\mathbf{t}=\mathbf{0}} \quad (1.44)$$

As for random variables, we have :

$$\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \quad \Leftrightarrow \quad M_{\mathbf{X}+\mathbf{Y}}(\mathbf{t}) = M_{\mathbf{X}}(\mathbf{t}) \times M_{\mathbf{Y}}(\mathbf{t}) \quad (1.45)$$

1.2.3 Ursell functions – or connected correlation functions

For random fields the cumulant-generating function is defined as for random variables :

$$g_{\mathbf{X}}(\mathbf{t}) = \log(M_{\mathbf{X}}(\mathbf{t})) \quad (1.46)$$

It generates so-called *connected correlation functions* or *Ursell functions* through :

$$\kappa_k(X_{i_1}, \dots, X_{i_k}) = \left. \frac{\partial^k g_{\mathbf{X}}(\mathbf{t})}{\partial t_{i_1} \dots \partial t_{i_k}} \right|_{\mathbf{t}=\mathbf{0}} \quad (1.47)$$

so that :

$$g_{\mathbf{X}}(\mathbf{t}) = \sum_{k=1}^{+\infty} \sum_{i_1 \dots i_k=1}^n \frac{t_{i_1} \dots t_{i_k}}{k!} \times \kappa_k(X_{i_1}, \dots, X_{i_k}) \quad (1.48)$$

The connected correlation functions can also be noted $\langle X_{i_1} \dots X_{i_k} \rangle_c$.

They are related to the simple correlation functions through formulas analog to Eq.1.16-1.21 linking cumulants and moments. In particular, both coincide up to order 3 and differ afterwards.

As for random variables, we have :

$$\mathbf{X} \perp\!\!\!\perp \mathbf{Y} \quad \Leftrightarrow \quad g_{\mathbf{X}+\mathbf{Y}}(\mathbf{t}) = g_{\mathbf{X}}(\mathbf{t}) + g_{\mathbf{Y}}(\mathbf{t}) \quad (1.49)$$

so that :

$$\kappa_k(\mathbf{X} + \mathbf{Y}) = \kappa_k(\mathbf{X}) + \kappa_k(\mathbf{Y}) \quad \forall k \quad (1.50)$$

This is an important property which gives conceptual and practical advantages to connected correlation functions over simple correlation functions.

A useful consequence is that if all X_i involved are either equal or independent (and at least two of them are different), then the connected correlation function is zero.

For example let's consider two independent random variables $Y \perp\!\!\!\perp Z$, and \mathbf{X} such that $X_1 = \dots = X_m = Y$ and $X_{m+1} = \dots = X_k = Z$. We define \mathbf{Y} and \mathbf{Z} by :

$$\begin{aligned} Y_1 = \dots = Y_m = Y & \quad \text{and} \quad Y_{m+1} = \dots = Y_k = 0 \\ Z_1 = \dots = Z_m = 0 & \quad \text{and} \quad Z_{m+1} = \dots = Z_k = Z \end{aligned}$$

where 0 is the random variable which returns zero everytime. We have $\mathbf{X} = \mathbf{Y} + \mathbf{Z}$ with $\mathbf{Y} \perp\!\!\!\perp \mathbf{Z}$. Hence

$$\begin{aligned} \kappa_k(X_1, \dots, X_k) &= \kappa_k(\mathbf{Y}) + \kappa_k(\mathbf{Z}) \\ &= 0 + 0 = 0 \end{aligned} \quad (1.51)$$

1.2.4 Gaussian random fields

A Gaussian random field \mathbf{X} is a continuous random field parametrised by its mean \mathbf{m} and an $n \times n$ matrix C :

$$p(\mathbf{X}) = \frac{1}{\sqrt{(2\pi)^n |\det C|}} \exp \left[-\frac{1}{2} {}^T (\mathbf{X}^* - \mathbf{m}^*) \cdot C^{-1} \cdot (\mathbf{X} - \mathbf{m}) \right] \quad (1.52)$$

with $*$ the complex conjugation, $\mathbf{X} = (X_i)_{i=1\dots n}$ considered as a vector, the dot \cdot is the matrix multiplication. Furthermore, C is the covariance matrix :

$$\forall i, j \quad C_{i,j} = \langle \tilde{X}_i \tilde{X}_j \rangle = \langle X_i X_j \rangle_c \quad (1.53)$$

or with vector notations $C = \langle \mathbf{X}^T \mathbf{X} \rangle_c$

Gaussian random fields appear frequently and have many applications in physics, signal processing etc. Out of their numerous properties, the following will be useful here :

- If the covariance matrix is diagonal, X_i are independent. In other words if two Gaussian random variables are uncorrelated, they are independent. This is the converse statement of the fact that independent random variables are uncorrelated, and generally does not hold if the variables are not Gaussian.
- $\kappa_1(\mathbf{X}) = \mathbf{m}$, $\kappa_2(\mathbf{X}) = C$, and all higher order connected correlation functions are zero. Moreover empirical measurement of the first and second order correlation functions is optimal, meaning that no further information on the parameters of the field can be extracted from the data (where data means empirical realisations of \mathbf{X}).

1.2.5 White noise

A white noise field is a collection of independent random variables, i.e. all X_i are independent. If the X_i are furthermore identically distributed (i.i.d.), the white noise is called *homogeneous*.

Sometimes the term white noise is used for Gaussian white noise, but we do not make this confusion here, as we will use white noises with different underlying p.d.f. later on.

Example : the instrumental noise in *Planck* maps is an inhomogeneous Gaussian white noise. Indeed each pixel is independent of the others (at least at first approximation), the instrumental noise is Gaussian, and some pixels are less noisy than others (because they have been observed more often due to the *Planck* scanning strategy).

Due to the properties described in Sect.1.2.3, the Ursell functions vanish for a white noise unless all X_i involved are equal :

$$\kappa_k(X_{i_1}, \dots, X_{i_k}) = \kappa_k(X_{i_1}) \delta_{i_1, \dots, i_k} \quad (1.54)$$

with :

$$\delta_{i_1, \dots, i_k} = \begin{cases} 1 & \text{if } i_1 = \dots = i_k \\ 0 & \text{otherwise} \end{cases} \quad (1.55)$$

and $\kappa_k(X_i)$ is the k-th order cumulant of X_i considered as a single random variable. Furthermore, if the white noise is homogeneous, $\kappa_k(X_i)$ does not depend on X_i .

1.2.6 Gram-Charlier expansion

As stated previously, full knowledge of the cumulants (/Ursell functions) of a random variable (/field) gives knowledge of the p.d.f. of the variable (/field), under some regularity condition. The Gram-Charlier expansion puts this idea in practice when only some of the cumulants are known. It is an expansion of the p.d.f. with respect to a fiducial p.d.f., in terms of the difference of the cumulants. Most often, and in this thesis in particular, the fiducial p.d.f. is the Gaussian one with same mean and variance. By inputting Eq.1.14 into Eq.1.8 via Eq.1.12 and taking the inverse Laplace transform, one gets for a random variable X :

$$p(x) = \exp \left[\sum_{n=3}^{+\infty} \frac{(-1)^n \kappa_n}{n!} \left(\frac{d}{dx} \right)^n \right] \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(x-\mu)^2}{2\sigma^2} \right] \quad (1.56)$$

Developing up to the fourth order we get :

$$p(x) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(x-\mu)^2}{2\sigma^2} \right] \times \left(1 + \frac{\kappa_3/\sigma^3}{3!} H_3 \left(\frac{x-\mu}{\sigma} \right) + \frac{\kappa_4/\sigma^4}{4!} H_4 \left(\frac{x-\mu}{\sigma} \right) + \mathcal{O}(\text{order } 5) \right) \quad (1.57)$$

with H_n the n -th order Hermite polynomial (Hermite 1864).

For a random field, computations become more cumbersome but we find up to the third order :

$$\begin{aligned} p(x_1, \dots, x_n) &= \exp \left[\sum_{n=3}^{+\infty} \frac{(-1)^n}{n!} \sum_{i_1 \dots i_n} \kappa_n(X_{i_1}, \dots, X_{i_n}) \frac{\partial^n}{\partial x_{i_1} \dots \partial x_{i_n}} \right] \\ &\quad \frac{1}{\sqrt{(2\pi)^n |\det C|}} \exp \left[-\frac{1}{2} {}^T (\mathbf{X} - \mathbf{m}) \cdot C^{-1} \cdot (\mathbf{X} - \mathbf{m}) \right] \quad (1.58) \\ &\approx \frac{1}{\sqrt{(2\pi)^n |\det C|}} \exp \left[-\frac{1}{2} {}^T (\mathbf{X} - \mathbf{m}) \cdot C^{-1} \cdot (\mathbf{X} - \mathbf{m}) \right] \left(1 - \right. \\ &\quad \frac{1}{3!} \sum_{i_1, i_2, i_3} \kappa_3(X_{i_1}, X_{i_2}, X_{i_3}) \left(\sum_a C_{i_1, i_2}^{-1} C_{i_3, a}^{-1} \tilde{x}_a + C_{i_1, i_3}^{-1} C_{i_2, a}^{-1} \tilde{x}_a + C_{i_2, i_3}^{-1} C_{i_1, a}^{-1} \tilde{x}_a \right. \\ &\quad \left. \left. + \sum_{a, b, c} C_{i_1, a}^{-1} C_{i_2, b}^{-1} C_{i_3, c}^{-1} \tilde{x}_a \tilde{x}_b \tilde{x}_c \right) \right) \quad (1.59) \end{aligned}$$

with $\tilde{x}_i = x_i - m_i$.

For concision, we can compact the Gaussian part in a single notation, and use Einstein's summation convention and the symmetrization operator $A_{[i_1 \dots i_n]} = (A_{i_1 \dots i_n} + \text{perm.}) / n!$. Eq.1.59 then shortens to :

$$p(x_1, \dots, x_n) \approx \text{Gauss}(\mathbf{X}) \left(1 - \frac{\kappa_3(X_i, X_j, X_k)}{3!} \left(3C_{[ij]a}^{-1} C_{k]a}^{-1} \tilde{x}_a - C_{i,a}^{-1} C_{j,b}^{-1} C_{k,c}^{-1} \tilde{x}_a \tilde{x}_b \tilde{x}_c \right) \right) \quad (1.60)$$

This shows that the third-order correlation function is the leading order indicator of non-Gaussianity of the field. Babich (2005) and Creminelli et al. (2006) have indeed

used the previous equation to devise an estimator for the primordial non-Gaussianity in the Cosmic Microwave Background, and we will use this expansion in Sect.2.2.4.3 to define the optimal estimator of the third order c.f. in harmonic space (the bispectrum).

In this chapter, I have thus described random variables/fields and how they can be characterised by cumulants/connected correlation functions. The measurement of these correlation functions hence allows to extract statistical information from realisations of the field. In particular, non-zero cumulants/connected correlation functions probe deviations from the Gaussian case.

Chapter 2

Some spherical analysis

Contents

2.1	Spherical harmonics	16
2.1.1	Definition	16
2.1.2	Useful properties	16
2.1.3	Relation with angular momentum and Gaunt coefficients . . .	17
2.2	Harmonic coefficients and polyspectra	18
2.2.1	Harmonic coefficients	18
2.2.2	Isotropic polyspectra	18
2.2.3	Relation with real-space correlation functions	20
2.2.4	Estimation of the power spectrum and bispectrum	21
2.2.4.1	Power spectrum	21
2.2.4.2	Bispectrum	23
2.2.4.3	Effect of masking on the bispectrum : variance	25
2.2.4.4	Effect of masking on the bispectrum : bias	27
2.2.4.5	Binned bispectrum coupling matrix evaluation	28
2.2.4.6	Bispectrum representation : a new parametrisation .	32
2.2.5	(Co)variance	32
2.2.5.1	Power spectrum	33
2.2.5.2	Bispectrum	34
2.2.5.3	Incomplete sky	37
2.3	Examples	37
2.3.1	Gaussian random field	37
2.3.2	White noise	38
2.3.2.1	Generating non-Gaussian fields	39
2.4	Relation with Fourier statistics	39
2.4.1	2D polyspectra	40
2.4.2	3D polyspectra	41

2.4.3	Plane wave expansion	42
2.4.4	Projection of the past light cone	42

In this chapter, we introduce the harmonic analysis used in cosmology. We will work principally on the sphere where *Planck* observations are set, but also in 3D where theoretical models are naturally set. Building up on the previous chapter, we describe correlation functions in harmonic space, i.e. polyspectra : their structure, their estimation, their error bars and their representation. We also illustrate these tools on examples of cosmological relevance.

2.1 Spherical harmonics

2.1.1 Definition

An element of the 2D sphere \mathbb{S}_2 is a unit vector \hat{n} . It can also be parametrised by the spherical coordinates angle $(\theta, \phi) \in [0, \pi] \times [0, 2\pi[$, where θ defines the latitude with respect to the North pole and ϕ defines the longitude.

The area element $d^2\hat{n}$ is the infinitesimal solid angle :

$$d^2\hat{n} = d\Omega = \sin\theta d\theta d\phi \quad (2.1)$$

measured in steradians (sr) with a total sky area of 4π steradians $\simeq 41\,000 \text{ deg}^2$.

The spherical harmonics $Y_{\ell m}$ are the eigenfunctions of the Laplacian on the sphere¹. As such, they define an orthogonal basis for the canonical scalar product². They are indexed by two integers : the multipole $\ell \in \mathbb{N}$ and the azimuthal parameter m with $-\ell \leq m \leq \ell$; alternatively they may be indexed by $s = \ell(\ell + 1) + m + 1$.

If we normalise them with respect to the canonical scalar product, the complex-valued spherical harmonics are :

$$Y_{\ell m}(\hat{n}) = \sqrt{\frac{2\ell + 1}{4\pi} \frac{(\ell - m)!}{(\ell + m)!}} P_{\ell}^m(\cos\theta) e^{im\phi} \quad (2.2)$$

with P_{ℓ}^m the associated Legendre polynomial (Abramowitz & Stegun 1964).

One can define real-valued spherical harmonics, by replacing $e^{im\phi}$ in Eq.2.2 with $\cos(m\phi)$ for $m > 0$ and $\sin(|m|\phi)$ for $m < 0$. However the cosmology convention is to work with complex-valued $Y_{\ell m}$.

2.1.2 Useful properties

The orthonormality condition reads :

$$\int d^2\hat{n} Y_{\ell m}(\hat{n}) Y_{\ell' m'}^*(\hat{n}) = \delta_{\ell\ell'} \delta_{mm'} \quad (2.3)$$

and the completeness (or closure) relation is :

$$\sum_{\ell m} Y_{\ell m}(\hat{n}) Y_{\ell m}^*(\hat{n}') = \delta^2(\hat{n} - \hat{n}') \quad (2.4)$$

¹with eigenvalue $\lambda_{\ell m} = \ell(\ell + 1) - m^2$.

²The Laplacian defines a positive-definite quadratic form through $\langle f, g \rangle = \int f \Delta g = \int \nabla f \cdot \nabla g = \int g \Delta f$. It can hence be diagonalised and the eigenfunctions form a complete set.

They express the fact that spherical harmonics form an orthonormal basis. Under complex conjugation we have :

$$Y_{\ell m}^* = (-1)^m Y_{\ell, -m} \quad (2.5)$$

The addition theorem allows to perform sums over the azimuthal parameter :

$$\sum_{m=-\ell}^{\ell} Y_{\ell m}(\hat{n}) Y_{\ell m}^*(\hat{n}') = \frac{2\ell+1}{4\pi} P_{\ell}(\hat{n} \cdot \hat{n}') \quad (2.6)$$

Notations :

In the following, I may note for concision i for a subscript (ℓ_i, m_i) , e.g. $Y_1 = Y_{\ell_1 m_1}$. A subscript i^* will correspond, through Eq.2.5, to a subscript $(\ell_i, -m_i)$ and a multiplicative factor $(-1)^{m_i}$.

Furthermore, I introduce the convention of using multiple indices to denote a product over it. E.g. for a quantity X , we have $X_{1\dots n} \equiv \prod_{i=1}^n X_i$.

For example, with these conventions Eq.2.3 implies :

$$\int d^2 \hat{n} Y_{12} = \delta_{12^*} \quad (2.7)$$

2.1.3 Relation with angular momentum and Gaunt coefficients

It can be shown (Condon & Shortley 1935) that the eigenstate of the (spinless) electron in the hydrogen atom admits a wavefunction of the form :

$$\psi(\vec{r}) = R_{n,\ell}(r) Y_{\ell m}(\theta, \phi) \quad (2.8)$$

with $n \in \mathbb{N}^*$ defining the energy state of the electron, ℓ defining its angular momentum L and m defining the projection of the angular momentum on the vertical axis L_z .

Spherical harmonics are the eigenvectors of the operators L^2 and L_z , with eigenvalue $\hbar \ell(\ell+1)$ and $\hbar m$ respectively, which hence gives physical interpretation to ℓ and m .

The sum of two angular momenta (ℓ_1, m_1) & (ℓ_2, m_2) yields a linear mixture of states with total angular momentum (ℓ_3, m_3) , the coefficients involved in this linear combination being the *Clebsch-Gordan* coefficients :

$$\langle \ell_1 m_1, \ell_2 m_2 | \ell_3 m_3 \rangle = (-1)^{\ell_1 - \ell_2 + m} \sqrt{2\ell_3 + 1} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \quad (2.9)$$

where the last term is the Wigner $3j$ symbol. It is zero unless the triangular conditions are met :

$$\ell_1 + \ell_2 + \ell_3 \text{ is even,} \quad (2.10)$$

$$|\ell_1 - \ell_2| \leq \ell_3 \leq \ell_1 + \ell_2, \quad (2.11)$$

$$m_1 + m_2 + m_3 = 0. \quad (2.12)$$

Correspondingly, the product of two spherical harmonics decomposes onto the harmonic basis as :

$$Y_{\ell_1 m_1} Y_{\ell_2 m_2} = \sum_{\ell_3 m_3} G_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} Y_{\ell_3 m_3}^* \quad (2.13)$$

where $G_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3} (\in \mathbb{R})$ is the Gaunt coefficient³ :

$$G_{123} = \int d^2 \hat{n} Y_{123}(n) \quad (2.14)$$

$$= \sqrt{\frac{(2\ell+1)_{123}}{4\pi}} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \quad (2.15)$$

Consequently, one can transform any product of spherical harmonics into a sum of single spherical harmonics. The integral of such a product is hence a combination of Gaunt coefficients :

$$\int d^2 \hat{n} Y_{1\dots n}(\hat{n}) = \sum_{\ell_1^d \dots \ell_{n-3}^d} \sum_{m_1^d \dots m_{n-3}^d} G_{1,2,1^d} \times G_{1^{d*},3,2^d} \cdots G_{(n-4)^{d*},n-2,(n-3)^d} \times G_{(n-3)^{d*},(n-1),n} \quad (2.16)$$

2.2 Harmonic coefficients and polyspectra

2.2.1 Harmonic coefficients

As the spherical harmonics form a basis, any function $T(\hat{n})$ on the sphere can be decomposed over it :

$$T(\hat{n}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{n}) \quad (2.17)$$

Through the orthonormality property Eq.2.3, the harmonic coefficients $a_{\ell m}$ can be expressed as :

$$a_{\ell m} = \int d^2 \hat{n} T(\hat{n}) Y_{\ell m}^*(\hat{n}) \quad (2.18)$$

and, if $T(\hat{n}) \in \mathbb{R}$, under complex conjugation we have :

$$a_{\ell m}^* = (-1)^m a_{\ell, -m} \quad (2.19)$$

Hence the set of coefficients $\{a_{\ell m} | m \geq 0\}$ provides a complete description of $T(\hat{n})$, with $a_{\ell m} \in \mathbb{C}$ for $m > 0$ and $a_{\ell 0} \in \mathbb{R}$.

2.2.2 Isotropic polyspectra

In the following we will consider statistically isotropic random fields on the sphere. This means that the probability law of the field is invariant under rotation and under parity :

$$\forall \mathcal{R} \in \mathcal{O}(3) \quad P \left[(T(\mathcal{R}\hat{n}_i))_{i=1 \dots n_{\text{pix}}} \right] = P \left[(T(\hat{n}_i))_{i=1 \dots n_{\text{pix}}} \right] \quad (2.20)$$

This implies that (connected) correlation functions are invariant under rotation, and in particular the 1-point c.f. is independent of the pixel : $\langle T(\hat{n}) \rangle = \langle T \rangle \forall \hat{n}$. In the following, we will assume that the fields have zero mean (or that it has been substracted), as this will simplify the expressions of real-space correlation functions.

³For simplicity, I note $G_{123} = G_{m_1 m_2 m_3}^{\ell_1 \ell_2 \ell_3}$ and $Y_i = Y_{\ell_i m_i}$ in the following.

Under the assumption of statistical isotropy, the harmonic coefficient average (1-point c.f.) are zero : $\langle a_{\ell m} \rangle = 0$.

Statistical isotropy also simplifies the harmonic 2-point c.f. :

$$\langle a_{\ell m} a_{\ell' m'}^* \rangle_c = C_\ell \delta_{\ell \ell'} \delta_{m m'} \quad (2.21)$$

In other terms, the covariance matrix is diagonal in harmonic space. This equation defines the angular power spectrum C_ℓ of the field, which possesses one degree of freedom (d.o.f.).

The harmonic 3-point c.f. is :

$$\langle a_1 a_2 a_3 \rangle_c = b_{\ell_1 \ell_2 \ell_3} G_{123} \quad (2.22)$$

which defines the angular bispectrum $b_{123} = b_{\ell_1 \ell_2 \ell_3}$ with three degrees of freedom⁴. The study of the bispectrum is the main subject of this thesis.

Generally, the harmonic n -point c.f. is :

$$\langle a_{1 \dots n} \rangle_c = \sum_{\ell_1^d \dots \ell_{n-3}^d} \mathcal{P}^{(n)}(\ell_1 \dots \ell_n, \ell_1^d \dots \ell_{n-3}^d) \times \mathcal{G}(\mathbf{1} \dots \mathbf{n}, \ell_1^d \dots \ell_{n-3}^d) \quad (2.23)$$

with :

$$\mathcal{G}(\mathbf{1} \dots \mathbf{n}, \ell_1^d \dots \ell_{n-3}^d) = \sum_{m_1^d \dots m_{n-3}^d} G_{1,2,1^d} G_{1^{d*},3,2^d} \dots G_{(n-4)^{d*},n-2,(n-3)^d} G_{(n-3)^{d*},(n-1),n} \quad (2.24)$$

which defines the angular polyspectrum $\mathcal{P}^{(n)}$ of order n with $2n - 3$ degrees of freedom. At orders $n \geq 4$ the degrees of freedom cannot be parametrised solely by the involved multipoles and additional *diagonal* degrees of freedom ($\ell_1^d \dots \ell_{n-3}^d$) are introduced. Geometrical interpretation of the polyspectra and of the diagonal d.o.f. is more obvious in the flat case (when we have a random field over \mathbb{R}^n instead of the sphere) and will hence be explained in Sect.2.4.1.

A particular case is when the polyspectrum does not vary with ($\ell_1^d \dots \ell_{n-3}^d$), it is then said *diagonal-independent* and Eq.2.23 takes the simpler form :

$$\langle a_{1 \dots n} \rangle_c = \mathcal{P}^{(n)}(\ell_1 \dots \ell_n) \mathcal{G}(\mathbf{1} \dots \mathbf{n}) \quad \text{with} \quad \mathcal{G}(\mathbf{1} \dots \mathbf{n}) = \int d^2 \hat{n} Y_{1 \dots n}(\hat{n}) \quad (2.25)$$

⁴ $b_{\ell_1 \ell_2 \ell_3}$ is sometimes called the reduced bispectrum in the literature, where the bispectrum is defined through $\langle a_1 a_2 a_3 \rangle_c = B_{\ell_1 \ell_2 \ell_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$. However the convention I have adopted is more natural as it is the analogous of the power spectrum at third order, it provides simpler expressions for theoretical bispectra, and reduces to the Fourier bispectrum in the flat-sky limit.

2.2.3 Relation with real-space correlation functions

The correlation functions in harmonic space are univoquely related to their real-space counterparts. Indeed we can write :

$$\begin{aligned}
\langle T(\hat{n}_1) \cdots T(\hat{n}_n) \rangle_c &= \sum_{\ell_i m_i} \langle a_{1 \dots n} \rangle_c [Y_i(\hat{n}_i)]_{i=1 \dots n} \\
&= \sum_{\ell_i \ell_j^d} \mathcal{P}^{(n)}(\ell_1 \cdots \ell_n, \ell_1^d \cdots \ell_{n-3}^d) \\
&\quad \times \sum_{m_i m_j^d} G_{1,2,1^d} G_{1^{d*},3,2^d} \cdots G_{(n-4)^{d*},n-2,(n-3)^d} G_{(n-3)^{d*},(n-1),n} [Y_i(\hat{n}_i)]_{i=1 \dots n} \\
&= \sum_{\ell_i \ell_j^d} \mathcal{P}^{(n)}(\ell_i, \ell_j^d) \times F_{\ell_i \ell_j^d}(\hat{n}_1, \cdots, \hat{n}_n)
\end{aligned} \tag{2.26}$$

where $F_{\ell_i \ell_j^d}$ is the contribution to the real-space n -point Ursell function of a given polyspectrum configuration. It is a rotationally invariant function for which there is no simple expression on the sphere to my knowledge, other than the definition implied by the previous lines. We can note nevertheless that, through the spherical harmonic closure relation Eq.2.4, we have :

$$\sum_{\ell_i \ell_j^d} F_{\ell_i \ell_j^d}(\hat{n}_1, \cdots, \hat{n}_n) = \delta^{(2)}(\hat{n}_1 - \hat{n}_2) \cdots \delta^{(2)}(\hat{n}_{n-1} - \hat{n}_n) \equiv \delta^{(2n)}(\hat{n}_1 = \cdots = \hat{n}_n) \tag{2.27}$$

At second order, $F_{\ell_i \ell_j^d}$ simplifies greatly :

$$\begin{aligned}
F_{\ell \ell'}(\hat{n}, \hat{n}') &= \sum_{m, m'} Y_{\ell m}(\hat{n}) Y_{\ell' m'}(\hat{n}') \int d^2 \hat{n}'' Y_{\ell m}(\hat{n}'') Y_{\ell' m'}(\hat{n}'') \\
&= \delta_{\ell \ell'} \sum_m Y_{\ell m}(\hat{n}) Y_{\ell m}^*(\hat{n}') = \delta_{\ell \ell'} \frac{2\ell + 1}{4\pi} P_\ell(\hat{n} \cdot \hat{n}')
\end{aligned} \tag{2.28}$$

Hence, with $\hat{n} \cdot \hat{n}' = \cos \theta$, we have :

$$C(\theta) \equiv \langle T(\hat{n}) T(\hat{n}') \rangle = \sum_\ell \frac{2\ell + 1}{4\pi} C_\ell P_\ell(\cos \theta) \tag{2.29}$$

In particular $\frac{\ell(2\ell+1)C_\ell}{4\pi}$ is the power (contribution to the variance) per logarithmic multipole bin.

At third order, the formula is more complicated :

$$\begin{aligned}
F_{\ell_1 \ell_2 \ell_3}(\hat{n}_1, \hat{n}_2, \hat{n}_3) &= \sum_{m_{123}} G_{123} Y_1(\hat{n}_1) Y_2(\hat{n}_2) Y_3(\hat{n}_3) \\
&= \int d^2 \hat{n} \sum_{m_{123}} Y_1^*(\hat{n}) Y_1(\hat{n}_1) Y_2^*(\hat{n}) Y_2(\hat{n}_2) Y_3^*(\hat{n}) Y_3(\hat{n}_3) \\
&= \frac{(2\ell + 1)_{123}}{(4\pi)^3} \int d^2 \hat{n} P_{\ell_1}(\hat{n} \cdot \hat{n}_1) P_{\ell_2}(\hat{n} \cdot \hat{n}_2) P_{\ell_3}(\hat{n} \cdot \hat{n}_3)
\end{aligned} \tag{2.30}$$

and will become even more complex at higher orders.

In the particular case where $\hat{n}_1 = \hat{n}_2 = \hat{n}_3$, $F_{\ell_1 \ell_2 \ell_3}$ reduces to :

$$F_{\ell_1 \ell_2 \ell_3}(\hat{n}_1 = \hat{n}_2 = \hat{n}_3) = \frac{(2\ell + 1)_{123}}{(4\pi)^2} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2 = \frac{N_{123}}{4\pi} \quad (2.31)$$

where I introduce preemptively $N_{123} = \frac{(2\ell+1)_{123}}{4\pi} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2$. N_{123} is the number of bispectrum modes for the triplet (ℓ_1, ℓ_2, ℓ_3) , and is the analogous to $2\ell + 1$ for the power spectrum.

Hence we have :

$$\kappa_3 = \langle T^3 \rangle = \sum_{\ell_{123}} \frac{N_{123}}{4\pi} b_{123} \quad (2.32)$$

where the sum runs over multipole triplets following the triangular condition, otherwise $N_{123} = 0$. We can conclude that $\frac{\ell_1 \ell_2 \ell_3 N_{123}}{4\pi} b_{123}$ is the contribution to the skewness per logarithmic multipole bin.

2.2.4 Estimation of the power spectrum and bispectrum

2.2.4.1 Power spectrum

From Eq.2.21 we have $\langle |a_{\ell m}|^2 \rangle = C_\ell$. We can thus introduce an unbiased estimator \hat{C}_ℓ of the power spectrum as :

$$\hat{C}_\ell = \frac{1}{2\ell + 1} \sum_m |a_{\ell m}|^2 \quad (2.33)$$

\hat{C}_ℓ can also be shown to possess the minimal variance among unbiased estimators⁵, i.e. it is optimal.

Numerically Eq.2.33 takes $\mathcal{O}(N_{\text{pix}}^{3/2})$ operations to compute from a given map. The harmonic transform to compute the $a_{\ell m}$ is numerically the bottleneck as it scales as $\mathcal{O}(N_{\text{pix}}^{3/2})$ while the sum over m scales as $\mathcal{O}(N_{\text{pix}})$.

Using the so-called scalemaps defined by Spergel & Goldberg (1999) :

$$T_\ell(\hat{n}) = \sum_m a_{\ell m} Y_{\ell m}(\hat{n}) \quad (2.34)$$

which contains a single multipole ℓ , we can formally rewrite Eq.2.33 as :

$$\hat{C}_\ell = \frac{1}{2\ell + 1} \int d^2 \hat{n} T_\ell(\hat{n})^2 \quad (2.35)$$

Numerically, this is an inefficient estimation method for the power spectrum, as the computation of all scalemaps up to $\ell_{\text{max}} \propto N_{\text{pix}}^{1/2}$ takes $\mathcal{O}(N_{\text{pix}}^2)$ operations.

A common complication is that we do not usually have a full-sky map of the signal to analyse, either because the instrument is not able to scan all the sky (e.g. ground-based telescope) or because some areas of the sky are too contaminated by other signals (typically the Milky Way emission). Hence we are analysing a masked map :

$$T_M(\hat{n}) = T(\hat{n}) \times M(\hat{n}) \quad (2.36)$$

⁵For example $|a_{\ell 0}|^2$ is an unbiased estimator of C_ℓ , with a greater variance than \hat{C}_ℓ

with $M(\hat{n})$ the mask function. $M(\hat{n})$ is typically 0 for masked pixels and 1 for observed pixels, but we may choose to smooth the edges (apodisation) to avoid “ringing”⁶ (see e.g. Ponthieu et al. 2011).

This multiplication in real space translates into a convolution in harmonic space, so that the harmonic coefficient of the masked map takes the form :

$$a_1(T_M) = \sum_{23} a_2(T) a_3(M) G_{1^*,2,3} \quad (2.37)$$

where $a_{\ell m}(T)$ and $a_{\ell m}(M)$ are the harmonic coefficients respectively of the true sky signal and of the mask.

Direct inversion of the whole convolution is not possible because information has been lost by masking. As the convolution is linear, correlation functions are convolved order by order i.e. the n^{th} order masked correlation function is a convolved version of the n^{th} order full-sky correlation function. However, there is no information leakage between different orders.

For example, the covariance matrix becomes non-diagonal (see e.g. Hivon et al. 2002) contrary to Eq.2.21, which was valid in full-sky; the best possible estimation of the power spectrum should use these off-diagonal terms. It has been however shown that an unbiased estimation of the power spectrum can be obtained considering the pseudo-spectrum of the masked map, and that the method is nearly optimal for high multipoles (e.g. Hivon et al. 2002). Indeed, if we define the pseudo-spectrum of the masked map :

$$\tilde{C}_\ell(T_M) = \frac{1}{2\ell+1} \sum_m |a_{\ell m}(T_M)|^2 \quad (2.38)$$

Some algebra shows that it is on average a convolved version of the true power spectrum :

$$\langle \tilde{C}_\ell(T_M) \rangle = \sum_{\ell'} \mathcal{M}_{\ell\ell'} C_{\ell'}(T) \quad (2.39)$$

where \mathcal{M} is the coupling matrix :

$$\mathcal{M}_{\ell_1\ell_2} = \frac{2\ell_2+1}{4\pi} \sum_3 |a_3(M)|^2 \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{pmatrix}^2 \quad (2.40)$$

If the available sky is sufficiently large, the coupling matrix can be inverted to yield an unbiased estimation of $C_\ell(T)$. For smaller fraction of observed sky, the pseudo-power spectrum can be binned and the binned coupling matrix becomes invertible if the binning is large enough (Hivon et al. 2002).

The observed sky fraction is given in term of the mask as :

$$f_{\text{SKY}} = \frac{1}{4\pi} \int d^2\hat{n} M(\hat{n})^2 \quad (2.41)$$

If the true power spectrum is varying slowly compared to the coupling matrix, it can be factored out of the sum in Eq.2.39 and we get :

$$\langle \tilde{C}_\ell(T_M) \rangle \approx C_\ell(T) \times \sum_{\ell'} \mathcal{M}_{\ell\ell'} = f_{\text{SKY}} C_\ell(T) \quad (2.42)$$

⁶Abrupt features in real space create long-tailed oscillations in harmonic space, e.g. the Fourier transform of a step function is a sinc function

This is the so-called f_{SKY} approximation and it is justified for large f_{SKY} and/or slow varying power spectrum (Komatsu et al. 2002). In particular, it is exact if the power spectrum is constant, e.g. for a white noise, see Sect.2.3.2.

The coupling matrix has $\ell_{\text{max}}^2 \propto N_{\text{pix}}$ elements, and after precomputation of Wigner symbols and of the harmonic transform of the mask, Eq.2.40 involves N_{pix} operations. Thus the whole computation scales as $\mathcal{O}(N_{\text{pix}}^2)$, which, although not fast, is numerically tractable for *Planck*-like data sets (50 millions pixels per map at full resolution).

I have implemented a code to compute the coupling matrix and estimate power spectra on masked maps. Indeed, I wanted a flexible code written in IDL language, so that it could interface easily with all my other codes. And I wanted to be able to use it simply and quickly, for example for my work described in Sect.5.2.3, without having to deal with the shortcomings of existing codes written in other programming languages. As an illustration, Fig.2.1 shows a mask used in *Planck* analysis and the corresponding coupling matrix for $\ell = 1 \cdots 50$.

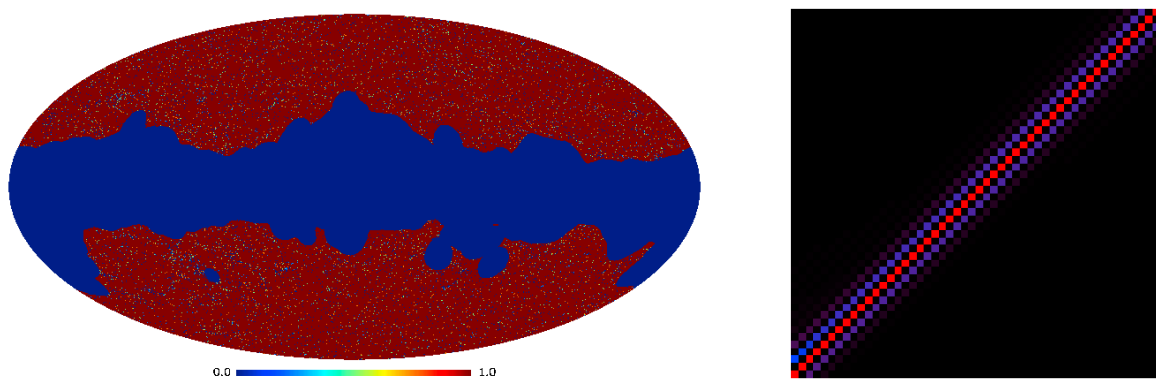


Figure 2.1: *Left* : galactic and point-source mask used in *Planck* Sunyaev-Zel'dovich analysis (Planck Collaboration XXI 2013). *Right* : power spectrum coupling matrix; the color code runs from black (zero) to red (maximum, here ~ 0.4). The coupling matrix is dominated by a band around the diagonal, as multipoles tend to couple to neighboring multipoles.

Let us note that even-odd matrix elements (e.g. $(\ell, \ell + 1)$ or $(\ell, \ell + 3)$) are small but non-zero. They would be exactly zero if the mask were parity-invariant. Indeed in the latter case $a_{\ell m}(M)$ is zero if ℓ is odd, and through Eq.2.40 and the parity condition for Wigner symbols ($\ell_1 + \ell_2 + \ell_3$ even), we see that $\mathcal{M}_{\ell_1 \ell_2}$ is zero if ℓ_1 and ℓ_2 have different parity.

2.2.4.2 Bispectrum

Similarly to the power spectrum case, we can define an unbiased estimator of the bispectrum by :

$$\hat{b}_{123} = \frac{1}{N_{123}} \sum_{m_{123}} G_{123} a_1 a_2 a_3 \quad (2.43)$$

Indeed the properties of Wigner symbols lead to :

$$\sum_{m_{12}} G_{124} G_{125} = \frac{(2\ell+1)_{12}}{4\pi} \begin{pmatrix} \ell_1 & \ell_2 & \ell_4 \\ 0 & 0 & 0 \end{pmatrix}^2 \delta_{45} \Rightarrow \sum_{m_{123}} G_{123} G_{123} = N_{123} \quad (2.44)$$

The estimator can also be shown to be optimal. It can be rewritten with scalemaps as (Bucher et al. 2010; Lacasa et al. 2012) :

$$\hat{b}_{123} = \frac{1}{N_{123}} \int d^2\hat{n} T_{\ell_1}(\hat{n}) T_{\ell_2}(\hat{n}) T_{\ell_3}(\hat{n}) \quad (2.45)$$

which, once the scalemaps are computed, takes $\mathcal{O}(N_{\text{pix}})$ operations for each multipole triplet. By contrast, Eq.2.43 runs over $\propto \ell_{\text{max}}^2 \propto N_{\text{pix}}$ elements but the Gaunt coefficients are costly to compute at high multipoles, so the scalemap method is computationally faster.

There are $\propto \ell_{\text{max}}^3 \propto N_{\text{pix}}^{3/2}$ triplets of multipoles, so the computation of a full bispectrum scales as $\mathcal{O}(N_{\text{pix}}^{5/2})$ operations. For *Planck*-like data sets with millions of pixels, this can be numerically challenging as it takes several CPU-days. The resulting measurement is also challenging to use as it consists of $\mathcal{O}(N_{\text{pix}}^{3/2})$ coefficients with low signal to noise ratio.

Several approaches to circumvent this problem have been proposed, and the one I have followed is the binned bispectrum method introduced by Bucher et al. (2010). Cosmological signals have in general a smooth scale dependence so that polyspectra are smooth functions of the multipoles. Hence one can bin multipoles together, with little loss of information unless the bin-size becomes too large (compared to the scale of variation of the bispectrum).

If we define the binned scalemaps as :

$$T_{\Delta\ell}(\hat{n}) = \sum_{\ell \in \Delta\ell, m} a_{\ell m} Y_{\ell m}(\hat{n}) \quad (2.46)$$

and the binned number of triplets as :

$$N_{\Delta\ell_1 \Delta\ell_2 \Delta\ell_3} = \sum_{\ell_i \in \Delta\ell_i} N_{123} \quad (2.47)$$

we can define the binned bispectrum estimator as (Lacasa et al. 2012):

$$\hat{b}_{\Delta\ell_1 \Delta\ell_2 \Delta\ell_3} = \frac{1}{N_{\Delta\ell_1 \Delta\ell_2 \Delta\ell_3}} \int d^2\hat{n} T_{\Delta\ell_1}(\hat{n}) T_{\Delta\ell_2}(\hat{n}) T_{\Delta\ell_3}(\hat{n}) \quad (2.48)$$

which estimates a weighted average of the bispectrum within the multipole bins. The binned bispectrum method drastically reduces the number of measured coefficients which goes as $\sim n_{\text{bin}}^3$. It also increases their signal-to-noise ratio (SNR) so that the resulting measurement can be visualised and interpreted. The drawbacks are a loss of statistical information and of fine variations of the bispectrum, which are both mitigated if the binning is not too large and/or if the bispectrum is smooth.

For the signals of interests in chapters 4 and 5, a binning as large as $\Delta\ell = 128$ can be used. Indeed for unclustered sources (a constant bispectrum, see chapter 4) binning does not lose any information, while for clustered sources I checked that the correlation between the binned bispectrum and the full bispectrum was higher than 94% for $\Delta\ell = 128$ and higher than 97% for $\Delta\ell = 64$.

2.2.4.3 Effect of masking on the bispectrum : variance

This section shows that a linear term is needed for the bispectrum estimator to lower its variance in the partial sky case. In fact, I will show that the linear term is needed for any estimation of a third order moment. This term however vanishes for the bispectrum in the isotropic case, which explains why it was not considered previously.

The computation of this section are based on the Gram-Charlier expansion (Sect.1.2.6) and in particular on Eq.1.60, which I will nevertheless need to develop to the next order $\mathcal{O}(\kappa_3^2)$. I will first illustrate the computation on the simpler case of the estimation of the third order cumulant of a random variable with independent samples, which will show the arising of the linear term. I will then generalise the result to the estimation of the bispectrum of a random field, in which case the algebra becomes cumbersome.

Let us first consider $X_1 \cdots X_n$, n random variables i.i.d. with zero mean and variance σ^2 . In that case the second and third order Ursell functions take the form :

$$C_{ij} = \sigma^2 \delta_{ij} \quad (2.49)$$

$$\kappa_3(X_i, X_j, X_k) = \kappa_3 \delta_{ijk} \quad (2.50)$$

Following Eq.1.58 and Eq.1.60, we have :

$$p(x_1 \cdots x_n | \kappa_3) = \text{Gauss}(\mathbf{X}) \times \left(1 + \underbrace{\frac{\kappa_3}{3!} \sum_i \frac{x_i^3}{\sigma^6} - \frac{3x_i}{\sigma^4}}_{= \frac{\kappa_3}{3! \sigma^3} \sum_i H_3\left(\frac{x_i}{\sigma}\right)} \right) + \frac{\kappa_3^2}{2 \cdot 3!^2} \sum_{ij} \frac{\partial^6}{\partial x_i^3 \partial x_j^3} \text{Gauss}(\mathbf{X}) \quad (2.51)$$

where H_n denotes the n -th order Hermite polynomial, and I find :

$$\frac{\partial^6}{\partial x_i^3 \partial x_j^3} \text{Gauss}(\mathbf{X}) = \text{Gauss}(\mathbf{X}) \times \frac{1}{\sigma^6} \left[H_3\left(\frac{x_i}{\sigma}\right) H_3\left(\frac{x_j}{\sigma}\right) - \left(H_6\left(\frac{x_i}{\sigma}\right) + H_3\left(\frac{x_i}{\sigma}\right)^2 \right) \delta_{ij} \right] \quad (2.52)$$

Following Baye's theorem (Bayes & Price 1763), the likelihood of the third order cumulant is $\mathcal{L}(\kappa_3) \propto p(x_1 \cdots x_n | \kappa_3)$. Hence a Maximum Likelihood Estimator (MLE) is found by solving :

$$\frac{\partial p(x_1 \cdots x_n | \kappa_3)}{\partial \kappa_3} = 0 \quad (2.53)$$

This is equivalent to :

$$\frac{1}{3!^2} \frac{\kappa_3}{\sigma^6} \sum_i H_6\left(\frac{x_i}{\sigma}\right) + H_3\left(\frac{x_i}{\sigma}\right)^2 - \frac{1}{3!^2} \frac{\kappa_3}{\sigma^6} \sum_{i \neq j} H_3\left(\frac{x_i}{\sigma}\right) H_3\left(\frac{x_j}{\sigma}\right) = \frac{1}{3! \sigma^3} \sum_i H_3\left(\frac{x_i}{\sigma}\right) \quad (2.54)$$

so that the MLE takes the simple form :

$$\hat{\kappa}_3 = \frac{3! \sigma^3 \sum_i H_3\left(\frac{x_i}{\sigma}\right)}{\sum_i H_6\left(\frac{x_i}{\sigma}\right) + H_3\left(\frac{x_i}{\sigma}\right)^2 - \sum_{i \neq j} H_3\left(\frac{x_i}{\sigma}\right) H_3\left(\frac{x_j}{\sigma}\right)} \quad (2.55)$$

Note that the Gram-Charlier expansion is valid if the p.d.f. is a perturbation of the Gaussian p.d.f., hence the computations developed above are valid in the weak non-Gaussianity (NG) limit. If the number n of samples is sufficiently large, following Babich

(2005), we can hence replace the denominator in Eq.2.55 by its Gaussian expectation value :

$$\left\langle \sum_i H_6 \left(\frac{x_i}{\sigma} \right) + H_3 \left(\frac{x_i}{\sigma} \right)^2 - \sum_{i \neq j} H_3 \left(\frac{x_i}{\sigma} \right) H_3 \left(\frac{x_j}{\sigma} \right) \right\rangle_{\text{Gauss}} = 0 + 6 - 0 = 6 \quad (2.56)$$

Hence the MLE takes the form :

$$\hat{\kappa}_3 = \frac{1}{n} \sum_i x_i^3 - 3\sigma^2 x_i \quad (2.57)$$

This estimator is obviously unbiased and contains an intuitive cubic part $\frac{1}{n} \sum_i x_i^3$, but also a linear part $\frac{1}{n} \sum_i 3\sigma^2 x_i$. Note that $X_i^3 - 3\sigma^2 X_i = X_i^3 - 3\langle X_i^2 \rangle X_i$ is the Wick product of X_i (Wick 1950).

The cubic term alone is an unbiased estimator : the linear term has zero expectation value. However the linear term allows the estimator to be optimal (in the weak NG limit). Indeed the Wick product of X_i is the cubic polynomial⁷ with the lowest variance in the Gaussian case (Donzelli et al. 2012).

To summarize, reducing the variance of the κ_3 estimator is achieved by replacing the cubic product in the intuitive estimator $\frac{1}{n} \sum_i x_i^3$ by the Wick product $x_i^3 - 3\sigma^2 x_i$.

Let us now turn to the estimation of the bispectrum of a spherical random field. The full-sky estimator introduced previously is :

$$\hat{b}_{123} = \frac{1}{N_{123}} \sum_{m_{123}} G_{123} a_1 a_2 a_3 \quad (2.58)$$

A straightforward generalisation inspired by the previous considerations is to replace the product of the harmonic coefficients by the corresponding Wick product :

$$\hat{b}_{123} = \frac{1}{N_{123}} \sum_{m_{123}} G_{123} (a_1 a_2 a_3 - \langle a_1 a_2 \rangle a_3 - \langle a_1 a_3 \rangle a_2 - \langle a_2 a_3 \rangle a_1) \quad (2.59)$$

However, if we start from the Gram-Charlier expansion for spherical harmonic coefficients and follow calculations similar to the κ_3 case above, we find that the situation is a bit more complex for the bispectrum :

- in the partial sky case, the covariance matrix is no longer diagonal, hence the C_{ij}^{-1} terms do not factor out of the sum in Eq.1.60. Hence the spherical harmonic coefficients have to be inverse-covariance filtered (also known as Wiener filtering). In theory, this just induces a different weighting of harmonic coefficients in the estimator Eq.2.59 and a corresponding change in the normalisation of the estimator. In practice, Wiener filtering maps is a computational challenge⁸, and Wiener filtered maps were not available for the signals I have studied. I have hence neglected the Wiener filtering and focused on a *pseudo-bispectrum* estimator, analogously to the pseudo-spectrum method.

⁷with unit leading coefficient

⁸e.g. the covariance matrix has $n_{\text{pix}} \times n_{\text{pix}}$ entries and is impossible to store nor invert for *Planck*-like data sets with $n_{\text{pix}} = \mathcal{O}(10^6)$.

- in the full-sky case, by maximising the likelihood with respect to b_{123} one obtains an equation for b_{123} hence defining straightforwardly the MLE. On the contrary in the partial sky case, one obtains an equation involving all bispectrum coefficients. Hence, one has to solve a linear system, and the bispectrum estimates are coupled, as different estimates may e.g. involve the same product $a_1 a_2 a_3$ (albeit with different weights). This situation, and the need to invert a coupling matrix, will be treated for the pseudo-bispectrum case in the following sections (Sect.2.2.4.4 & 2.2.4.5).

Thus, we see that Eq.2.59 defines a pseudo-bispectrum estimator which needs to be debiased (see Sect.2.2.4.4 & 2.2.4.5) and is not optimal (in all cases, the optimality would have been guaranteed only in the weak NG limit). The situation is analogous to the pseudo-spectrum method, which does not provide an optimal estimator but is however useful because it is fast and near-optimal at high multipoles (Efstathiou 2004). The pseudo-bispectrum estimator Eq.2.59 can be rewritten with scale maps as :

$$\hat{b}_{123} = \frac{1}{N_{123}} \int d^2\hat{n} (T_{\ell_1} T_{\ell_2} T_{\ell_3} - \langle T_{\ell_1} T_{\ell_2} \rangle T_{\ell_3} - \langle T_{\ell_1} T_{\ell_3} \rangle T_{\ell_2} - \langle T_{\ell_2} T_{\ell_3} \rangle T_{\ell_1}) \quad (2.60)$$

where $T_\ell = T_\ell(\hat{n})$ for clarity.

Note that the average $\langle \rangle$ only involve 2-point correlations, hence they are in practice computed through average over Gaussian simulations. Moreover, in the full-sky case $\langle T_\ell(\hat{n}) T_{\ell'}(\hat{n}) \rangle$ is either zero (if $\ell \neq \ell'$) or a constant (if $\ell = \ell'$), so that the linear term vanishes in the isotropic case.

Note also that Eq.2.60 is straightforwardly generalised when multipoles are binned. Furthermore, binning multipoles decreases the number of simulations needed for convergence of the linear term, because binning induces an averaging and reduces the coupling between bins.

2.2.4.4 Effect of masking on the bispectrum : bias

As for the power spectrum, the estimation of the bispectrum is biased when part of the map is masked for any reason. I have worked on this subject, following the same idea as for the power spectrum. Indeed, if we compute the pseudo bispectrum of the masked map⁹ :

$$\tilde{b}_{123}(T_M) = \frac{1}{N_{123}} \sum_{m_{123}} G_{123} a_1(T_M) a_2(T_M) a_3(T_M) \quad (2.61)$$

The average $\tilde{b}_{123}(T_M)$ is a convolved version of the true bispectrum :

$$\langle \tilde{b}_{\ell_{123}}(T_M) \rangle = \sum_{\ell'_{123}} \mathcal{M}_{\ell_{123}, \ell'_{123}} b_{\ell'_{123}}(T) \quad (2.62)$$

⁹Considering only the cubic term, as the linear term has zero average

I derived several possible formulae for the bispectrum coupling matrix :

$$\begin{aligned} \mathcal{M}_{\ell_{123}, \ell'_{123}} &= \frac{1}{N_{123}} \int d^2 \hat{n}_{123} M(\hat{n}_1) M(\hat{n}_2) M(\hat{n}_3) \\ &\quad \times F_{\ell_1 \ell_2 \ell_3}(\hat{n}_1, \hat{n}_2, \hat{n}_3) F_{\ell'_1 \ell'_2 \ell'_3}(\hat{n}_1, \hat{n}_2, \hat{n}_3) \end{aligned} \quad (2.63)$$

$$= \frac{1}{N_{123}} \sum_{\substack{m_{123} \\ m'_{123} \\ 1''2''3''}} G_{123} G_{1'2'3'} (G_{i^*i'i''} a_{i''}(M))_{i=123} \quad (2.64)$$

$$= \frac{1}{N_{123}} \sum_{\substack{m_{123} \\ m'_{123}}} G_{123} G_{1'2'3'} W_{1,1'} W_{2,2'} W_{3,3'} \quad (2.65)$$

with $W_{i,i'} = W_{\ell_i m_i, \ell'_i m'_i}$, and :

$$W_{\ell m, \ell' m'} = \int d^2 \hat{n} M(\hat{n}) Y_{\ell m}(\hat{n}) Y_{\ell' m'}^*(\hat{n}) \quad (2.66)$$

which measures the non-orthonormality of spherical harmonics on the masked sky. To the best of my knowledge, the equations for the bispectrum coupling matrix do not simplify any further on the sphere. In all cases, the numerical computation of a full bispectrum coupling matrix is unconceivable as it has $\mathcal{O}(N_{\text{pix}}^3)$ elements.

As for the power spectrum, the equations simplify when the bispectrum varies slowly compared to the coupling matrix. Indeed, we then have :

$$\langle \tilde{b}_{\ell_{123}}(T_M) \rangle \simeq b_{\ell_{123}}(T) \times \sum_{\ell'_{123}} \mathcal{M}_{\ell_{123}, \ell'_{123}} \quad (2.67)$$

and through Eq.2.27-2.31-2.63 we have :

$$\begin{aligned} \sum_{\ell'_{123}} \mathcal{M}_{\ell_{123}, \ell'_{123}} &= \frac{1}{N_{123}} \int d^2 \hat{n} M^3(\hat{n}) F_{\ell_1 \ell_2 \ell_3}(\hat{n}, \hat{n}, \hat{n}) \\ &= \frac{1}{4\pi} \int d^2 \hat{n} M^3(\hat{n}) = f_{\text{SKY}} \end{aligned}$$

Hence we find again the f_{SKY} approximation $\langle \tilde{b}_{\ell_{123}}(T_M) \rangle \simeq f_{\text{SKY}} \times b_{\ell_{123}}(T)$.

2.2.4.5 Binned bispectrum coupling matrix evaluation through simulations

For a binned bispectrum estimator, we would not need a full bispectrum coupling matrix but a binned coupling matrix, with $\mathcal{O}(n_{\text{bin}}^6)$ elements instead of $\mathcal{O}(\ell_{\text{max}}^6)$. Such matrix sizes are manageable, and hence I have embarked in the project to find a method to compute such a binned coupling matrix. The purpose is to have a tool allowing me to debias binned bispectrum estimations on a masked sky. As the direct computation of (binned) Eq.2.63-2.65 is not possible, I have designed an alternative method : the evaluation of the coupling matrix through simulations.

Let us note n_{config} the number of binned bispectrum configurations, and let us consider bispectra as vectors with length n_{config} . Then Eq.2.62 can be rewritten as :

$$\mathbf{b}^{\text{obs}} = \mathcal{M} \cdot \mathbf{b}^{\text{true}} + N \quad (2.68)$$

where N is a noise vector and \mathcal{M} is an $n_{\text{config}} \times n_{\text{config}}$ matrix. If we have n_{sim} simulations with different input bispectra, noting :

$$\mathcal{B}_o = (\mathbf{b}_1^{\text{obs}} \cdots \mathbf{b}_{n_{\text{sim}}}^{\text{obs}}) \quad \text{and} \quad \mathcal{B}_t = (\mathbf{b}_1^{\text{true}} \cdots \mathbf{b}_{n_{\text{sim}}}^{\text{true}}) \quad (2.69)$$

we have :

$$\mathcal{B}_o = \mathcal{M} \cdot \mathcal{B}_t + N \quad (2.70)$$

Thus the idea is to generate maps with different bispectra such that $\mathbf{b}_1^{\text{true}} \cdots \mathbf{b}_{n_{\text{sim}}}^{\text{true}}$ form a spanning set of the bispectrum vector space, mask the maps and measure the corresponding masked pseudo-bispectra.

Then from the input \mathcal{B}_t and the observable \mathcal{B}_o , I should construct an estimator $\hat{\mathcal{M}}$ of \mathcal{M} . To this end, I choose an estimator of the form $\hat{\mathcal{M}} = \mathcal{B}_o \cdot W$, where W is an $n_{\text{sim}} \times n_{\text{config}}$ matrix. I am looking for W such that :

$$\left\{ \begin{array}{l} \mathcal{B}_t W = \text{Id} \\ \left\langle \text{Tr}({}^T \hat{\mathcal{M}} \hat{\mathcal{M}}) \right\rangle \text{ minimal} \end{array} \right. \quad (2.71)$$

These equations express the fact that I search for an unbiased estimator with minimal variance. The solution to this constrained minimization problem can be obtained through the method of Lagrange multipliers. After some algebra I find :

$$\hat{\mathcal{M}} = \mathcal{B}_o R_o^{-1} {}^T \mathcal{B}_t (\mathcal{B}_t R_o^{-1} {}^T \mathcal{B}_t)^{-1} \quad \text{with} \quad R_o = \langle {}^T \mathcal{B}_o \mathcal{B}_o \rangle \quad (2.72)$$

In the case where $n_{\text{sim}} = n_{\text{config}}$, all matrices are square ones and I get :

$$\hat{\mathcal{M}} = \mathcal{B}_o \mathcal{B}_t^{-1} \quad (2.73)$$

which is the obvious solution of Eq.2.70 with square matrices and negligible noise. I have adopted this approach, simpler than with rectangular matrices, and decided to reduce the noise in the estimation by repeating the procedure and averaging over iterations. That is to say I first use $n_{\text{sim}} = n_{\text{config}}$ simulations and obtain a first estimation $\hat{\mathcal{M}}_1$, I repeat ... and my final estimation is $\hat{\mathcal{M}}_f = \sum_i \hat{\mathcal{M}}_i / n_{\text{iter}}$.

The main question is how to generate simulations such that their bispectra form a basis, and that they have a high signal-to-noise ratio. Indeed, we want the smallest possible noise in the bispectra so that we have the smallest possible noise in the estimation of \mathcal{M} .

A bispectrum is a function of (ℓ_1, ℓ_2, ℓ_3) invariant under their permutations. As such it decomposes over symmetric polynomials :

$$\sigma_1 = \ell_1 + \ell_2 + \ell_3 \quad (2.74)$$

$$\sigma_2 = \ell_1 \ell_2 + \ell_1 \ell_3 + \ell_2 \ell_3 \quad (2.75)$$

$$\sigma_3 = \ell_1 \ell_2 \ell_3 \quad (2.76)$$

So the aim is to generate any polynomial of $(\sigma_1, \sigma_2, \sigma_3)$, i.e.

$$\begin{aligned} &1 \\ &\sigma_1, \sigma_2, \sigma_3 \\ &\sigma_1^2, \sigma_1 \sigma_2, \sigma_1 \sigma_3, \sigma_2^2, \sigma_2 \sigma_3, \sigma_3^2 \\ &\dots \end{aligned}$$

As will be shown in Sect.2.3.2, I am able to quickly generate maps with a theoretical bispectrum of the form :

$$b_{123} = \alpha_{\ell_1} \alpha_{\ell_2} \alpha_{\ell_3} \quad (2.77)$$

with α_ℓ an arbitrary function of ℓ , and with each bispectrum coefficient having a high SNR.

At zeroth order we get easily $b_{123} = 1$ by taking $\alpha_\ell = 1$.

At first order we can recover the three symmetric polynomials using $\alpha_\ell = \ell - \tilde{\ell}_{(0)}$, $\alpha_\ell = \ell - \tilde{\ell}_{(1)}$ and $\alpha_\ell = \ell - \tilde{\ell}_{(2)}$ with $\tilde{\ell}_{(0)} \neq \tilde{\ell}_{(1)} \neq \tilde{\ell}_{(2)}$ arbitrary constants. Indeed :

$$(\ell_1 - \tilde{\ell}_{(i)})(\ell_2 - \tilde{\ell}_{(i)})(\ell_3 - \tilde{\ell}_{(i)}) = \tilde{\ell}_{(i)}^2 \sigma_1 - \tilde{\ell}_{(i)} \sigma_2 + \sigma_3 \quad (2.78)$$

so we can invert¹⁰ the linear system between the three bispectra and the three symmetric polynomials.

It can also be seen that the six symmetric second-order polynomials can be recovered with linear combinations of the six bispectra determined by $\alpha_\ell = (\ell - \tilde{\ell}_{(i)})(\ell - \tilde{\ell}_{(j)})$ with $(i, j) \in \llbracket 0, 1, 2 \rrbracket$. This is also true at higher orders.

For simplicity we may note an α_ℓ with the indices of the $\tilde{\ell}_{(i)}$ involved, e.g. $(\ell - \tilde{\ell}_{(0)})(\ell - \tilde{\ell}_{(1)})$ can be noted 01. Then the 10 polynomials at third order can be noted in the form :

000, 100, 110, 111, 200, 210, 211, 220, 221, 222

and all α_ℓ can be ordered through lexicographic order.

I have developed this computation, then checked that the resulting bispectra indeed form a basis and that the coupling matrix is the identity for simulations without mask. However, the problem that I have encountered is that the matrices involved in the computation are poorly-conditioned¹¹ so that the noise gets amplified. For example with the α_ℓ described previously and choosing $\tilde{\ell}_{(0)} = \ell_{\min}$ $\tilde{\ell}_{(1)} = \ell_{\text{mean}}$ $\tilde{\ell}_{(2)} = \ell_{\max}$ ¹², with multipole bins $\Delta\ell = 64$ up to $\ell_{\max} = 256$, the condition number of \mathcal{B}_t is equal to 2×10^5 . Thus the recovered coupling matrices are quite noisy. Several iterations are needed to estimate this noise and decrease it through averaging. I have tried several changes to decrease the condition number, e.g. $\alpha_\ell = \sin(\ell - \tilde{\ell}_{(i)})$ or (many) other functional forms. For a fixed binning scheme $\Delta\ell = 64$ up to $\ell_{\max} = 256$, I have been able to reduce significantly the condition number, to $\mathcal{O}(100)$, and to obtain correct estimations of coupling matrices. For example, Fig.2.2 shows the matrix obtained with a 80% galactic mask and this binning scheme, giving 4 multipole bins and 13 bispectrum configurations. Namely the center of the multipole bins are 32, 96, 160, 224 and the bispectrum configurations are : (32,32,32), (32,96,96), (32,160,160), (32,224,224), (96,96,96), (96,96,160), (96,160,160), (96,160,224), (96,224,224), (160,160,160), (160,160,224), (160,224,224) and (224,224,224) ; these are the 13 entries of the matrix.

The matrix is dominated by its diagonal, and the off-diagonal terms are in fact dominated by noise. The diagonal terms are close to f_{SKY} , as the binning size is large and we have a large sky fraction. Indeed, we saw for the power spectrum that the mask mostly

¹⁰The determinant is that of Vandermonde and is non-zero iff $\tilde{\ell}_{(0)} \neq \tilde{\ell}_{(1)} \neq \tilde{\ell}_{(2)}$.

¹¹The condition number for a matrix M and a norm $\|\cdot\|$ is $\text{cond}(M) = \|M\| \|M^{-1}\|$. It quantifies the sensitivity of the numerical inversion of the matrix. Indeed for a matrix with a large condition number, small errors on the matrix coefficients lead to large errors on the coefficient of the invert. In our case, the cosmic variance of the simulations produces noise in the \mathcal{B}_0 coefficients.

¹²So that the α_ℓ are as different as possible.

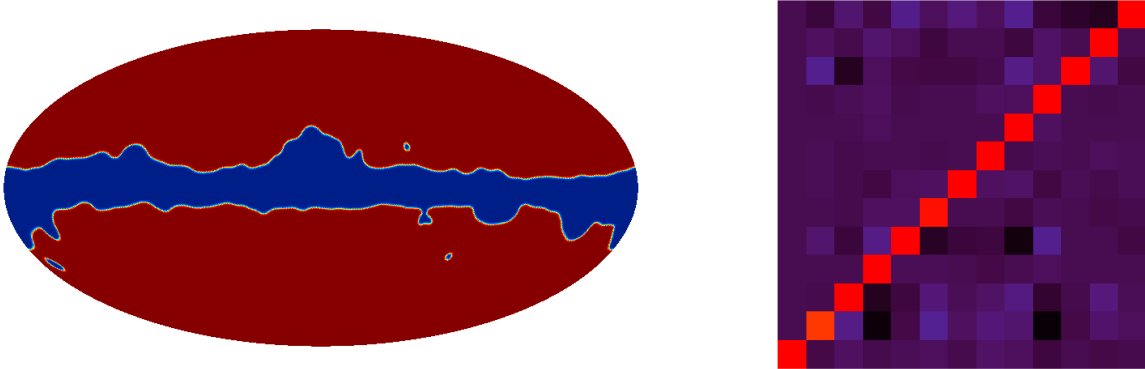


Figure 2.2: *Left* : 80% apodized galactic mask. *Right* : Corresponding binned bispectrum coupling matrix for $\Delta\ell = 64$ $\ell_{\max} = 256$ and 100 iterations ; the entries of the matrix are the bispectrum configurations.

couples neighboring multipoles, in the case of Fig.2.2 the correlation length is smaller than the bin size so that the coupling between different bins is negligible.

However, this method to compute numerically the bispectrum coupling matrix is not computationally conceivable for higher resolutions and/or numbers of multipole bins. Indeed, at higher resolution each bispectrum evaluation becomes slower. Moreover, if we increase the number of multipole bins, the number of configurations increases even more ($n_{\text{config}} \propto n_{\text{bin}}^3$) which increases the number of required simulations for each iteration. Furthermore, for binning schemes others than the previously described one, I have not been able to obtain condition number low enough (≤ 100) to make computations numerically possible.

Although I have not been able to compute coupling matrices at high enough resolution to debias the bispectrum measurement which will be described in Chapt.5, there is already valuable information in the results of Fig.2.2. Indeed, we see that the coupling between different bins for this bin size is quite negligible until $\ell_{\max} = 256$. Furthermore, we have seen for the power spectrum that the correlation length between multipoles does not increase at high multipoles, it even tends to decrease. Hence we can reasonably assume that for $\Delta\ell = 64$, the correlation between bins will remain negligible at high resolution. Thus, for bispectra without steep variations, we expect the f_{SKY} approximation to produce a correct debiasing of binned bispectra for this bin size. I will describe in Chapt.5 how I dealt with the problem for measurement on *Planck* data.

Finally, one of the perspective of this project is to find novel ways to reduce the condition number. Indeed if we had condition number of order $\mathcal{O}(1)$, only a few iterations would be necessary per matrix. Thus the computation would become possible at higher resolution in some CPU-days/weeks, up to $\ell_{\max} = 1024$ $N_{\text{side}} = 512$, which is a resolution high enough for interesting studies of the foregrounds that will be described later on (see Chapt.4 & 5). Optimising the α_ℓ to form a better basis, or other methods to generate simulations with a given bispectrum (e.g. Brown 2013) are other possible tracks to solve the problem.

2.2.4.6 Bispectrum representation : a new parametrisation

Once we have an optimal estimate of the bispectrum, we need a tool to visualise this complex three-dimensional quantity. Several ways of visualising the angular bispectrum have been proposed in the literature, e.g. isosurfaces in the (ℓ_1, ℓ_2, ℓ_3) 3D space by Fergusson & Liguori (2010), or slices of constant perimeter in the orthogonal transverse coordinate $(\ell_{\perp a}, \ell_{\perp b})$ space by Bucher et al. (2010).

We know that $b_{\ell_1 \ell_2 \ell_3}$ is invariant under permutations of ℓ_1 , ℓ_2 and ℓ_3 , i.e. it is a function of the shape and size of the triangle (ℓ_1, ℓ_2, ℓ_3) only. However the aforementioned visualisations do not account for this property and hence plot up to six times the same information. In Lacasa et al. (2012) I have proposed a parametrisation invariant under permutation of ℓ_1 , ℓ_2 , and ℓ_3 , as this will avoid redundancy of information and allow convenient visualisation and interpretation of data. Let us first denote $\overline{(\ell_1, \ell_2, \ell_3)}$ the equivalence class of the triplet under permutations.

The elementary symmetric polynomials ensure the invariance under permutations:

- $\sigma_1 = \ell_1 + \ell_2 + \ell_3$
- $\sigma_2 = \ell_1 \ell_2 + \ell_1 \ell_3 + \ell_2 \ell_3$
- $\sigma_3 = \ell_1 \ell_2 \ell_3$

Through Cardan's formula, there is a one-to-one correspondence between $\overline{(\ell_1, \ell_2, \ell_3)}$ and the triplet $(\sigma_1, \sigma_2, \sigma_3)$, as the multipoles are defined by the roots of the polynomial $X^3 - \sigma_1 X^2 + \sigma_2 X - \sigma_3$.

I further define the scale-invariant parameters $\tilde{\sigma}_2 = 12\sigma_2/\sigma_1^2 - 3$ and $\tilde{\sigma}_3 = 27\sigma_3/\sigma_1^3$ with coefficients chosen so that $\tilde{\sigma}_2$ and $\tilde{\sigma}_3$ vary in the range $[0,1]$. As illustrated in the upper panel of Fig. 2.3, this parametrisation does not allow us to discriminate efficiently between the different triangles.

I have introduced the parameters noted (P, F, S) :

- $P = \sigma_1$ (perimeter)
- $F = 32(\tilde{\sigma}_2 - \tilde{\sigma}_3)/3 + 1$
- $S = \tilde{\sigma}_3$

which provide a clearer distinction of the triangles as is illustrated in the bottom panel of Fig. 2.3.

I have adopted this parametrisation to plot bispectra in the following way : defining perimeter bins between the minimal ($P_{\min} = 3\ell_{\min}$) and the maximal ($P_{\max} = 3\ell_{\max}$) perimeter, I plot all the triangles in a given perimeter bin in the (F, S) space and code the value of the bispectrum with a color. In this thesis, bispectra will be plotted either with this new parametrisation or in some particular configurations (e.g., equilateral depending on $\ell = \ell_1 = \ell_2 = \ell_3$).

2.2.5 (Co)variance

A measurement is useless without the associated error bar. Here I derive the (co)variance of a power spectrum or bispectrum measurement in the full-sky case (see Sefusatti et al.

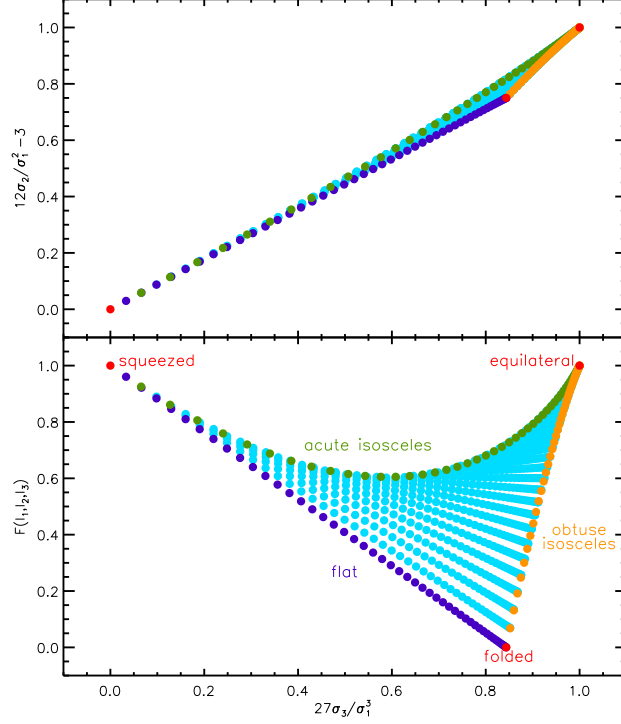


Figure 2.3: *Upper panel* : triangles of constant perimeter, P , in the parametrisation defined by the normalised symmetric polynomials. *Bottom panel* : same triangles in the new parameter space (F, S) proposed by Lacasa et al. (2012) and defined in the text.

(2006) for a derivation in the 3D Fourier case and Joachimi et al. (2009) for the 2D flat sky limit). The partial sky case is more complex, and I describe an approximation in this case in Sect.2.2.5.3.

2.2.5.1 Power spectrum

The power spectrum estimator \hat{C}_ℓ is :

$$\hat{C}_\ell = \frac{1}{2\ell + 1} \sum_m a_{\ell m} a_{\ell m}^* \quad (2.79)$$

Hence we have :

$$\begin{aligned} \langle \hat{C}_\ell \hat{C}_{\ell'} \rangle &= \frac{1}{(2\ell + 1)(2\ell' + 1)} \sum_{m, m'} \langle a_{\ell m} a_{\ell m}^* a_{\ell' m'} a_{\ell' m'}^* \rangle \\ &= \frac{1}{(2\ell + 1)(2\ell' + 1)} \sum_{m, m'} [\langle a_{\ell m} a_{\ell m}^* \rangle \langle a_{\ell' m'} a_{\ell' m'}^* \rangle + \langle a_{\ell m} a_{\ell' m'} \rangle \langle a_{\ell m}^* a_{\ell' m'}^* \rangle \\ &\quad + \langle a_{\ell m} a_{\ell' m'}^* \rangle \langle a_{\ell' m'} a_{\ell m} \rangle + \langle a_{\ell m} a_{\ell m}^* a_{\ell' m'} a_{\ell' m'}^* \rangle_c] \\ &= C_\ell C_{\ell'} + \delta_{\ell\ell'} \frac{2C_\ell C_{\ell'}}{2\ell + 1} + \frac{1}{(2\ell + 1)(2\ell' + 1)} \sum_{m, m'} (-1)^{m+m'} \\ &\quad \times \sum_{\ell^d} \mathcal{P}^{(4)}(\ell, \ell, \ell', \ell', \ell^d) \sum_{m^d} G_{\ell\ell\ell^d}^{m, -m, m^d} \times (-1)^{m^d} G_{\ell^d\ell'\ell'}^{-m^d, m', -m'} \end{aligned}$$

$$\begin{aligned}
&= C_\ell C_{\ell'} + \delta_{\ell\ell'} \frac{2C_\ell C_{\ell'}}{2\ell+1} + \frac{1}{(2\ell+1)(2\ell'+1)} \sum_{m,m',\ell^d} \mathcal{P}^{(4)}(\ell, \ell, \ell', \ell', \ell^d) \\
&\times \sum_{m^d} \frac{1}{\sqrt{4\pi}} \delta_{\ell^d,0} \delta_{m^d,0} \frac{1}{\sqrt{4\pi}} \delta_{\ell^d,0} \delta_{m^d,0} \\
&= C_\ell C_{\ell'} + \delta_{\ell\ell'} \frac{2C_\ell C_{\ell'}}{2\ell+1} + \frac{\mathcal{P}^{(4)}(\ell, \ell, \ell', \ell', \ell^d=0)}{4\pi}
\end{aligned}$$

so that :

$$\text{Cov}(\hat{C}_\ell, \hat{C}_{\ell'}) = \frac{2C_\ell C_{\ell'}}{2\ell+1} \delta_{\ell\ell'} + \frac{\mathcal{P}^{(4)}(\ell, \ell, \ell', \ell', \ell^d=0)}{4\pi} \quad (2.80)$$

The first term is the Gaussian contribution and the second term the trispectrum contribution. Let us note that the 4π denominator allows the trispectrum contribution to have the same unit ($[\mathcal{P}^{(4)}] = K^4 \text{sr}^3$) as the covariance ($[C_\ell^2] = K^4 \text{sr}^2$). We can also note that the covariance matrix is diagonal in the Gaussian case, where the coefficients $a_{\ell m}$ and thus the power spectrum estimations are independent.

2.2.5.2 Bispectrum

The bispectrum estimator \hat{b}_{123} is :

$$\hat{b}_{123} = \frac{1}{N_{123}} \sum_{m_{123}} G_{123} a_{123} \quad (2.81)$$

Hence we have :

$$\langle \hat{b}_{123} \hat{b}_{456} \rangle = \frac{1}{N_{123} N_{456}} \sum_{m_{1\dots 6}} G_{123} G_{456} \langle a_{1\dots 6} \rangle \quad (2.82)$$

Through the relations between correlation functions and Ursell functions, $\langle a_{1\dots 6} \rangle$ splits into four terms :

$$\langle a_{1\dots 6} \rangle = \langle a_{1\dots 6} \rangle_{2 \times 2 \times 2} + \langle a_{1\dots 6} \rangle_{3 \times 3} + \langle a_{1\dots 6} \rangle_{4 \times 2} + \langle a_{1\dots 6} \rangle_c \quad (2.83)$$

The first term corresponds to the case when $a_{\ell m}$ are grouped by pairs; it contains 15 terms of the form $\langle a_{12} \rangle \langle a_{34} \rangle \langle a_{56} \rangle$. The second term corresponds to the case when $a_{\ell m}$ are grouped by triplets; it contains 10 terms of the form $\langle a_{123} \rangle \langle a_{456} \rangle$. The third term corresponds to the case when $a_{\ell m}$ are grouped into a 4-uplet (trispectrum) and a pair; it contains 15 terms of the form $\langle a_{1234} \rangle_c \langle a_{56} \rangle$. The last term is the sixth-order polyspectrum contribution.

The derivation is relatively long, so that I will proceed term by term and give the results without the derivation steps. For the first term ($2 \times 2 \times 2$), three pairs have to be formed, which leads to two types of contributions :

- 9 contributions where at last one of the pair belongs to one of the bispectrum triplet (e.g. (12)(34)(56) : (12) belongs to the triplet (123)). These contributions are null except if the monopole is considered (which is not the case in this thesis). Example : (12)(34)(56) $\rightarrow \delta \langle \hat{b}_{123} \hat{b}_{456} \rangle_{2 \times 2 \times 2} = 0$ except if $\ell_3 = \ell_4 = 0$.

- 6 contributions where only multipoles from different triplets are paired (e.g. (14)(25)(36)). These contributions are non-zero when the multipoles paired are identical.

Example : $(14)(25)(36) \rightarrow \delta \left\langle \hat{b}_{123} \hat{b}_{456} \right\rangle_{2 \times 2 \times 2} \stackrel{C_{\ell_1} C_{\ell_2} C_{\ell_3}}{=} \frac{C_{\ell_1} C_{\ell_2} C_{\ell_3}}{N_{123}} \delta_{\ell_1 \ell_4} \delta_{\ell_2 \ell_5} \delta_{\ell_3 \ell_6}$

All this can be shortened into :

$$\left\langle \hat{b}_{123}^2 \right\rangle_{2 \times 2 \times 2} = \frac{C_{\ell_1} C_{\ell_2} C_{\ell_3}}{N_{123}} \times \begin{cases} 6 & \text{equilateral triangle} \\ 2 & \text{isosceles triangle} \\ 1 & \text{general triangle} \end{cases} \quad (2.84)$$

and $\left\langle \hat{b}_{123} \hat{b}_{456} \right\rangle_{2 \times 2 \times 2} = 0$ if ℓ_{123} and ℓ_{456} are different triangles.

For the second term (3×3) , two triplets have to be formed, which leads to two types of contributions :

- 1 contribution where the triplets corresponds to the original bispectrum triplets :

$$(123)(456) \rightarrow \delta \left\langle \hat{b}_{123} \hat{b}_{456} \right\rangle_{3 \times 3} = b_{123} b_{456}$$

This contribution cancels the term $\left\langle \hat{b}_{123} \right\rangle \left\langle \hat{b}_{456} \right\rangle$ in the definition of the covariance.

- 9 other contributions where two multipoles of the original triplets have been permuted (e.g. (124)(356) : 3 and 4 have been permuted).

Example : $(124)(356) \rightarrow \delta \left\langle \hat{b}_{123} \hat{b}_{456} \right\rangle_{3 \times 3} = \frac{b_{124} b_{356}}{2\ell_3 + 1} \delta_{\ell_3 \ell_4}$

Note that for a bispectrum of the form $b_{123} = \alpha_{\ell_1} \alpha_{\ell_2} \alpha_{\ell_3}$, the numerators of the 9 contributions (e.g. $b_{124} b_{356}$) are all equal to $b_{123} b_{456}$.

In a single expression, we have :

$$\begin{aligned} \text{Cov}(\hat{b}_{123}, \hat{b}_{456})_{3 \times 3} &= \frac{b_{156} b_{234}}{2\ell_1 + 1} \delta_{\ell_1 \ell_4} + \frac{b_{146} b_{235}}{2\ell_1 + 1} \delta_{\ell_1 \ell_5} + \frac{b_{145} b_{236}}{2\ell_1 + 1} \delta_{\ell_1 \ell_6} \\ &+ \frac{b_{134} b_{256}}{2\ell_2 + 1} \delta_{\ell_2 \ell_4} + \frac{b_{135} b_{246}}{2\ell_2 + 1} \delta_{\ell_2 \ell_5} + \frac{b_{136} b_{245}}{2\ell_2 + 1} \delta_{\ell_2 \ell_6} \\ &+ \frac{b_{124} b_{356}}{2\ell_3 + 1} \delta_{\ell_3 \ell_4} + \frac{b_{125} b_{346}}{2\ell_3 + 1} \delta_{\ell_3 \ell_5} + \frac{b_{126} b_{345}}{2\ell_3 + 1} \delta_{\ell_3 \ell_6} \end{aligned} \quad (2.85)$$

This contribution to the covariance matrix has off-diagonal terms.

For the third term (4×2) , a pair and a quadruplet have to be formed which leads to two types of contributions :

- 6 contributions where the pair belongs to one of the bispectrum triplet (e.g. (12)(3456) : (12) belongs to (123)). These contributions are null except if the monopole is considered (which is not the case in this thesis).

Example : $(12)(3456) \rightarrow \delta \left\langle \hat{b}_{123} \hat{b}_{456} \right\rangle_{4 \times 2} = 0$ except if $\ell_3 = 0$

- 9 contributions where multipoles paired are from different bispectrum triplets.

Example : $(14)(2356) \rightarrow \delta \left\langle \hat{b}_{123} \hat{b}_{456} \right\rangle_{4 \times 2} = \frac{C_{\ell_1} \mathcal{P}^{(4)}(\ell_2, \ell_3, \ell_5, \ell_6, \ell^d = \ell_1)}{2\ell_1 + 1} \delta_{\ell_1 \ell_4}$

In a single formula, we have :

$$\begin{aligned} \text{Cov}(\hat{b}_{123}, \hat{b}_{456})_{4 \times 2} = & \frac{C_{\ell_1} \mathcal{P}^{(4)}(\ell_{2356}, \ell_1)}{2\ell_1 + 1} \delta_{\ell_1 \ell_4} + \frac{C_{\ell_1} \mathcal{P}^{(4)}(\ell_{2346}, \ell_1)}{2\ell_1 + 1} \delta_{\ell_1 \ell_5} + \frac{C_{\ell_1} \mathcal{P}^{(4)}(\ell_{2345}, \ell_1)}{2\ell_1 + 1} \delta_{\ell_1 \ell_6} \\ & + \frac{C_{\ell_2} \mathcal{P}^{(4)}(\ell_{1356}, \ell_2)}{2\ell_2 + 1} \delta_{\ell_2 \ell_4} + \frac{C_{\ell_2} \mathcal{P}^{(4)}(\ell_{1346}, \ell_2)}{2\ell_2 + 1} \delta_{\ell_2 \ell_5} + \frac{C_{\ell_2} \mathcal{P}^{(4)}(\ell_{1345}, \ell_2)}{2\ell_2 + 1} \delta_{\ell_2 \ell_6} \\ & + \frac{C_{\ell_3} \mathcal{P}^{(4)}(\ell_{1256}, \ell_3)}{2\ell_3 + 1} \delta_{\ell_3 \ell_4} + \frac{C_{\ell_3} \mathcal{P}^{(4)}(\ell_{1246}, \ell_3)}{2\ell_3 + 1} \delta_{\ell_3 \ell_5} + \frac{C_{\ell_3} \mathcal{P}^{(4)}(\ell_{1245}, \ell_3)}{2\ell_3 + 1} \delta_{\ell_3 \ell_6} \end{aligned} \quad (2.86)$$

This term of the covariance matrix contributes in the same off-diagonal entries as the 3×3 term.

The last term (from the 6-th order polyspectrum) has fortunately only one contribution :

$$(123456) \rightarrow \delta \left\langle \hat{b}_{123} \hat{b}_{456} \right\rangle_c = \frac{\mathcal{P}^{(6)}(\ell_{1\dots 6}, \ell_1^d = \ell_3, \ell_2^d = 0, \ell_3^d = \ell_4)}{4\pi} \quad (2.87)$$

I do not group all terms into a single equation, and I will only need to refer to Eq.2.84-2.85-2.86-2.87 individually in the following. Note however that, similarly to the power spectrum, the covariance matrix is diagonal in the Gaussian case. However bispectrum measurements are then not independent but only linearly uncorrelated : indeed they mix up different multipoles and e.g., $\hat{b}_{\ell_1 \ell_2 \ell_3}$ and $\hat{b}_{\ell_1 \ell'_2 \ell'_3}$ are obviously not independent. In the non-Gaussian case, the bispectrum covariance matrix is clearly non-diagonal.

All these terms can be understood simply and graphically in the Fourier case where, as explained in Sect.2.4.1, n^{th} order polyspectra correspond to a 2D polygon with n sides. All terms of the bispectrum covariance correspond to possibilities of forming polygons with 6 vectors (1,2,3,4,5,6) given that (1,2,3) and (4,5,6) respectively form triangles. Fig.2.4 shows diagrams representative of the different terms of the bispectrum covariance matrix.

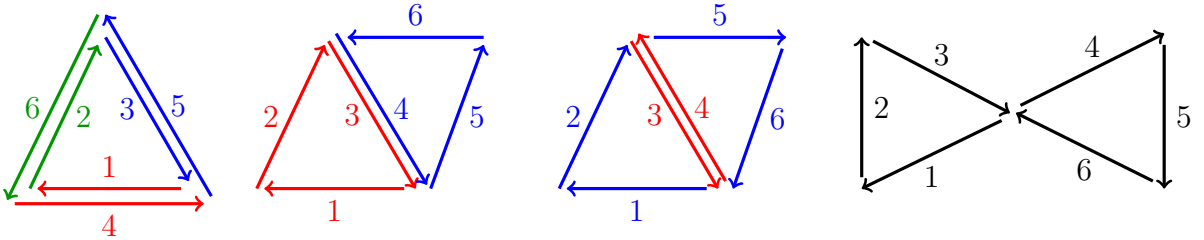


Figure 2.4: Diagrams corresponding to the different terms of the bispectrum covariance. From left to right $2 \times 2 \times 2$ (double-circulation triangle), 3×3 (parallel kite), 4×2 (opposite kite) and 6 (butterfly).

In the same color are multipoles which are averaged together to form a polyspectrum. For example in the first diagram (1,4) (2,6) and (3,5) are averaged together respectively to form a power spectrum, yielding $C_{\ell_1} \delta_{14^*}$ $C_{\ell_2} \delta_{26^*}$ and $C_{\ell_3} \delta_{35^*}$ respectively, so that the total contribution of the diagram is $\frac{C_{\ell_1} C_{\ell_2} C_{\ell_3}}{N_{123}} \delta_{\ell_1 \ell_4} \delta_{\ell_2 \ell_6} \delta_{\ell_3 \ell_5}$.

2.2.5.3 Incomplete sky

In the incomplete sky case, we have seen that the estimation of a polyspectrum becomes complicated. The covariance matrix of this estimation is even more complicated; for example at the power spectrum level the covariance matrix may be computed with simulations or approximated analytically, but a full analytical computation is numerically virtually impossible (see e.g. Efstathiou 2004).

However, the situation becomes tractable if the polyspectrum varies slowly compared to the coupling matrix. Indeed, we have seen in the previous sections that the polyspectrum estimation is then biased by a factor f_{SKY} . The covariance of this estimation is a particular case of the polyspectrum at double order, i.e., the covariance of the n^{th} order polyspectrum involves the $2n^{\text{th}}$ order polyspectrum. This covariance is hence also multiplied by a factor f_{SKY} .

As we want an unbiased estimate of the polyspectrum, we need to divide the partial-sky measurement by f_{SKY} . For this unbiased estimate, we then have :

$$\text{COV}_{\text{partial-sky}} = \text{COV}_{\text{full-sky}} / f_{\text{SKY}} \quad (2.88)$$

This implies that the error bars scale as $1/\sqrt{n_{\text{pix}}}$ (at fixed pixel size), i.e. going down with the number of observations.

This result has been derived in the case where the polyspectrum varies slowly compared to the mask coupling effect. In particular it is exact for flat polyspectra, e.g., that of a white noise. However it can be extended to larger masks with binned estimation, if the bin size renders the coupling between different bins negligible. The latter property is verified if the binned coupling matrix is mostly diagonal (Hivon et al. 2002). For example this is the case in Fig.2.2. In the following, we will assume that the bin size is large enough to be in this situation.

2.3 Examples

2.3.1 Gaussian random field

An isotropic Gaussian random field is entirely described by its mean $\langle T \rangle$ and its power spectrum C_ℓ . Furthermore all the statistical information is contained in these two correlation functions, i.e., the measurement of higher order correlation functions does not provide any improvement on the constraint of $\langle T \rangle$ and C_ℓ .

Ordering the harmonic coefficients in lexicographic order, the probability of a map reads :

$$P(\mathbf{a} = (a_{\ell m})) = \frac{1}{\sqrt{(2\pi)^N \det \mathbf{C}}} \exp \left[{}^T \bar{\mathbf{a}} \mathbf{C}^{-1} \mathbf{a} \right] \quad (2.89)$$

$$\text{with } \mathbf{C} = \begin{pmatrix} C_1 & 0 & \cdots & & \cdots & 0 \\ 0 & C_1 & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & & & \\ & & & C_\ell & & \\ & & & & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & C_\ell & 0 \\ 0 & \cdots & & \cdots & 0 & \ddots \end{pmatrix} \quad (2.90)$$

$\underbrace{\hspace{10em}}_{\ell + 1 \text{ terms}}$

where C_ℓ is the power spectrum of the field and all higher order polyspectra are zero. Thus $a_{\ell m}$ are random complex numbers (real for $m = 0$) with variance C_ℓ , and importantly they are independent (except of course $a_{\ell m}$ and $a_{\ell, -m}$). The statistical information is contained in the moduli of the harmonic coefficients while their phases are independent and distributed randomly and uniformly in $[0, 2\pi[$.

The important example in cosmology of a Gaussian random field on the sphere is the Cosmic Microwave Background (CMB), which will be described in Sect.3.3. Another example is instrumental noise, which is often Gaussian to a good approximation. The noise may however not be homogeneous, as is the case for *Planck*, and as it is often uncorrelated between different pixels it is better characterised in pixel space than in harmonic space.

2.3.2 White noise

For a white noise field all pixels are independent. If the white noise is homogeneous all pixels are drawn from the same underlying distribution and in particular the field is isotropic. This is for example the case of the emission of unclustered extragalactic point-sources. As explained in Sect.1.2.5, in this case the pixel space correlation function takes the form :

$$\langle T(\hat{n}_1) \cdots T(\hat{n}_n) \rangle_c = \kappa_n(T) \delta_{\hat{n}_1, \dots, \hat{n}_n} \quad (2.91)$$

with a Kronecker symbol being one iff $\hat{n}_1 \cdots \hat{n}_n$ are the same pixel. A harmonic transform then gives :

$$\langle a_1 \cdots a_n \rangle_c = \kappa_n(T) \Omega_{\text{pix}}^{n-1} \int d^2\hat{n} Y_{1\dots n}(\hat{n}) \quad (2.92)$$

where Ω_{pix} is the pixel area in steradian.

Thus a homogeneous white noise generates a diagonal-independent polyspectrum which is furthermore constant :

$$\mathcal{P}^{(n)}(\ell_1 \cdots \ell_n) = \kappa_n(T) \Omega_{\text{pix}}^{n-1} \quad (2.93)$$

In particular, the power spectrum is $\sigma^2 \Omega_{\text{pix}}$ (a well-known formula for instrumental noise), and the bispectrum is $\kappa_3 \Omega_{\text{pix}}^2$. This also makes the polyspectrum unit obvious, e.g. if the field $T(\hat{n})$ is measured in Kelvins, the n-order polyspectrum has unit $\text{K}^n \text{sr}^{n-1}$.

The important example in cosmology of a white noise field on the sphere is a distribution of unclustered point-sources, such as extragalactic radio point-sources described in Sect.4.3.1. Instrumental noise may also be whiten in addition to its Gaussianity.

2.3.2.1 Generating non-Gaussian fields

Using white noises is a convenient way to generate highly non-Gaussian maps with a given bispectrum form. The first step is to generate a white noise such that its bispectrum has the highest possible signal-to-noise ratio. For this purpose, I have used a Poisson law (see Sect.1.1.4.2) as the underlying distribution, since it is easily produced numerically by random number generators, and since the Poisson parameter λ is easily tunable to obtain a high SNR. Indeed, compiling the results of Sect.1.1.4.2 and Sect.2.2.5, for a Poisson white noise the SNR of a bispectrum coefficient is (up to some constants A, B, C, D) :

$$\begin{aligned} (\text{SNR})^2 &\propto \frac{(\lambda \Omega_{\text{pix}}^2)^2}{A (\lambda \Omega_{\text{pix}})^3 + B (\lambda \Omega_{\text{pix}}^2)^2 + C \lambda \Omega_{\text{pix}} \lambda \Omega_{\text{pix}}^3 + D \lambda \Omega_{\text{pix}}^5} \\ &= \frac{1}{B + C + A \lambda / \Omega_{\text{pix}} + D \Omega_{\text{pix}} / \lambda} \end{aligned} \quad (2.94)$$

The SNR is maximal for $\lambda \sim \Omega_{\text{pix}} \propto 1/n_{\text{pix}}$. So a given bispectrum configuration has the highest signal-to-noise ratio for a Poisson distribution with one source per map on average. In reality I have runned simulations with an average thousand sources per map, so that the covariance matrix shall be closer to diagonal, and thus different bispectrum configurations shall be few correlated.

These highly non-Gaussian simulations can then be processed in the following way :

- compute the harmonic transform of the white noise $\rightarrow a_{\ell m}^{\text{white}}$
- multiply the harmonic coefficient by an arbitrary function α_ℓ : $a_{\ell m}^{\text{sim}} = \alpha_\ell a_{\ell m}^{\text{white}}$
- compute the inverse harmonic transform : $T_{\text{sim}}(\hat{n}) = \sum_{\ell m} a_{\ell m}^{\text{sim}} Y_{\ell m}(\hat{n})$

We then easily see that the final simulation has a bispectrum :

$$b_{123} = \alpha_{\ell_1} \alpha_{\ell_2} \alpha_{\ell_3} \lambda \Omega_{\text{pix}}^3 \quad (2.95)$$

with the same (high) SNR as the original white noises, and that this bispectrum fulfills the requirement of Eq.2.77 for the coupling matrix estimation¹³.

2.4 Relation with Fourier statistics

This section describes the statistical tools introduced previously, in the case of a field $f(\mathbf{x})$ over \mathbb{R}^d ($d = 2, 3$), instead of over the sphere. Indeed, sufficiently small patches of the sphere may be analysed as flat fields, with a correspondence between the Fourier modes and the spherical harmonics : the so-called ‘flat-sky’ limit (see e.g. Bernardeau et al. 2011, for a formalisation). Furthermore modelisations of cosmological field (e.g. the galaxy density, see Sect.3.4.3) are naturally set in 3D, hence I will describe 3D statistics and how they project onto the sphere, when observations cannot probe the radial distance. In the following, the Fourier transform of the field will be noted $a_{\mathbf{k}} = \int d^d \mathbf{x} f(\mathbf{x}) e^{i\mathbf{k} \cdot \mathbf{x}}$.

¹³The $\lambda \Omega_{\text{pix}}^3$ can be incorporated in a redefinition of α_ℓ .

2.4.1 2D polyspectra

For a 2D field, the polyspectrum of order n is defined through :

$$\langle a_{\mathbf{k}_1} \cdots a_{\mathbf{k}_n} \rangle_c = (2\pi)^2 \delta^{(2)}(\mathbf{k}_1 + \cdots + \mathbf{k}_n) \mathcal{P}^{(n)}(\mathbf{k}_1, \cdots, \mathbf{k}_n) \quad (2.96)$$

where $(\mathbf{k}_1, \cdots, \mathbf{k}_n)$ form a polygon in 2D. Figure 2.5 shows the corresponding polygon for the second (power spectrum), third (bispectrum), fourth (trispectrum) orders and the general case. At orders ≥ 4 , the polygon cannot be parametrised only by the length of its side, but some diagonals need to be fixed. This is why diagonal degrees of freedom appear in the polyspectrum. The choice of the fixed diagonals is somewhat arbitrary, for example for the trispectrum one can either fix $|\mathbf{k}_1 + \mathbf{k}_2| = |\mathbf{k}_3 + \mathbf{k}_4|$ or fix $|\mathbf{k}_1 + \mathbf{k}_4| = |\mathbf{k}_2 + \mathbf{k}_3|$. In analogy with my choice on the sphere, I here take a choice of diagonals inherited from the orders of the \mathbf{k}_i and shown in Fig.2.5 (i.e. $\mathbf{k}_1^d = \mathbf{k}_1 + \mathbf{k}_2$, $\mathbf{k}_2^d = \mathbf{k}_1 - \mathbf{k}_3$ etc). Expanding the dirac $\delta^{(2)}(\mathbf{k}_1 + \cdots + \mathbf{k}_n)$ by decomposing the polygon in its sub-triangles, Eq.2.96 takes the following form :

$$\begin{aligned} \langle a_{\mathbf{k}_1} \cdots a_{\mathbf{k}_n} \rangle_c &= \int \frac{d^2 \mathbf{k}_1^d}{(2\pi)^2} \cdots \frac{d^2 \mathbf{k}_{n-3}^d}{(2\pi)^2} \mathcal{P}^{(n)}(k_1, \cdots, k_n, k_1^d, \cdots, k_{n-3}^d) \\ &\quad (2\pi)^2 \delta^{(2)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_1^d) \times (2\pi)^2 \delta^{(2)}(-\mathbf{k}_1^d + \mathbf{k}_3 + \mathbf{k}_2^d) \\ &\quad \times \cdots \times (2\pi)^2 \delta^{(2)}(-\mathbf{k}_{n-3}^d + \mathbf{k}_{n-1} + \mathbf{k}_n) \end{aligned} \quad (2.97)$$

Let us note in particular that this is exactly analogous to the definition of the polyspectrum on the sphere (Eq.2.23-2.24) with the replacements :

$$\begin{aligned} \sum_{\ell^d m^d} &\rightarrow \int \frac{d^2 \mathbf{k}^d}{(2\pi)^2} \\ G_{123} &\rightarrow (2\pi)^2 \delta^{(2)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \end{aligned}$$

The last line reflects intuition as both terms are the integral of the basis functions :

$$G_{123} = \int d^2 \hat{n} Y_{123}(\hat{n}) \quad \text{and} \quad (2\pi)^2 \delta^{(2)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) = \int d^2 \mathbf{x} e^{i(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \cdot \mathbf{x}} \quad (2.98)$$

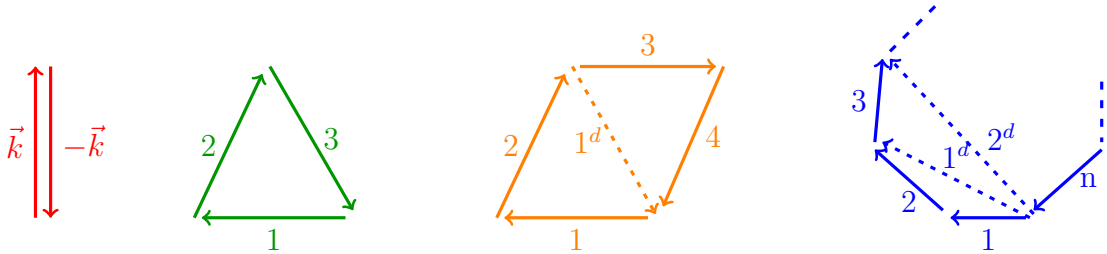


Figure 2.5: Diagrams corresponding to the 2D polyspectrum. From left to right : orders 2, 3, 4 and n .

2.4.2 3D polyspectra

For a 3D field, the polyspectrum of order n is defined through :

$$\langle a_{\mathbf{k}_1} \cdots a_{\mathbf{k}_n} \rangle_c = (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \cdots + \mathbf{k}_n) \mathcal{P}^{(n)}(\mathbf{k}_1, \cdots, \mathbf{k}_n) \quad (2.99)$$

where $(\mathbf{k}_1, \cdots, \mathbf{k}_n)$ forms a polygon in 3D so that the polyspectrum has $3n - 6$ degrees of freedom¹⁴. A power spectrum still has one d.o.f. : $\mathcal{P}^{(2)}(\mathbf{k}_1, \mathbf{k}_2) = P(k = |\mathbf{k}_1|)$ and can be visualised as two opposite vectors. A bispectrum still has three d.o.f. : $\mathcal{P}^{(3)}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = B(k_1, k_2, k_3)$ and can be visualised as a triangle.

To parametrise the polyspectrum d.o.f., we must choose a triangulation of the 3D polygon, then the polyspectrum can be parametrised by the length of the n sides and $2n - 6$ diagonals. We will index with $g = 1 \cdots 2n - 5$ the triangles of the minimal triangulation chosen, and note $(\mathbf{k}_1(g), \mathbf{k}_2(g), \mathbf{k}_3(g))$ the vectors of this triangle. This vectors $\mathbf{k}_i(g)$ are either a $\pm k_i$ or a $\pm k_j^d$.

The Fourier n -point c.f. then takes the form :

$$\begin{aligned} \langle a_{\mathbf{k}_1} \cdots a_{\mathbf{k}_n} \rangle_c &= \int \frac{d^3 \mathbf{k}_1^d}{(2\pi)^3} \cdots \frac{d^3 \mathbf{k}_{2n-6}^d}{(2\pi)^3} \mathcal{P}^{(n)}(k_1, \cdots, k_n, k_1^d, \cdots, k_{2n-6}^d) \\ &\times \prod_g (2\pi)^3 \delta^{(3)}(\mathbf{k}_1(g) + \mathbf{k}_2(g) + \mathbf{k}_3(g)) \end{aligned} \quad (2.100)$$

Examples :

Fig.2.6 shows the polygons and their diagonals at third and fourth orders.

For the bispectrum, the polygon is a triangle; it is flat, so the number of d.o.f. coincides with the 2D bispectrum, and it has no diagonals. The triangulation is trivial : $g = 1$ and $\mathbf{k}_1(g) = \mathbf{k}_1, \mathbf{k}_2(g) = \mathbf{k}_2, \mathbf{k}_3(g) = \mathbf{k}_3$. In this case Eq.2.100 takes the form :

$$\langle a_{\mathbf{k}_1} \cdots a_{\mathbf{k}_3} \rangle_c = B(k_1, k_2, k_3) (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \quad (2.101)$$

For the trispectrum, the polygon is a tetrahedron and there are two diagonals : $\mathbf{k}_1^d = -\mathbf{k}_1 - \mathbf{k}_2 = \mathbf{k}_3 + \mathbf{k}_4$ and $\mathbf{k}_2^d = -\mathbf{k}_1 - \mathbf{k}_4 = \mathbf{k}_2 + \mathbf{k}_3$.

There are hence 6 d.o.f. which already diverges from the 2D case (5 d.o.f.). The tetrahedron has four triangular faces but only three need to be fixed, so there is some arbitrary decision in the triangulation. If we choose as triangulation $(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_1^d), (-\mathbf{k}_1^d, \mathbf{k}_3, \mathbf{k}_4), (\mathbf{k}_1, \mathbf{k}_2^d, \mathbf{k}_4)$, Eq.2.100 takes the form :

$$\begin{aligned} \langle a_{\mathbf{k}_1} \cdots a_{\mathbf{k}_4} \rangle_c &= \int \frac{d^3 \mathbf{k}_1^d}{(2\pi)^3} \frac{d^3 \mathbf{k}_2^d}{(2\pi)^3} \mathcal{P}^{(4)}(k_1, \cdots, k_4, k_1^d, k_2^d) (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_1^d) \\ &\times (2\pi)^3 \delta^{(3)}(-\mathbf{k}_1^d + \mathbf{k}_3 + \mathbf{k}_4) (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2^d + \mathbf{k}_4) \end{aligned} \quad (2.102)$$

¹⁴There are three d.o.f per vector, and six d.o.f. are discarded by isotropy, where six is the number of d.o.f. of a solid in 3D. Comparatively, a polyspectrum has $2n - 3$ d.o.f. in 2D, where three is the number of d.o.f. of a solid in 2D.

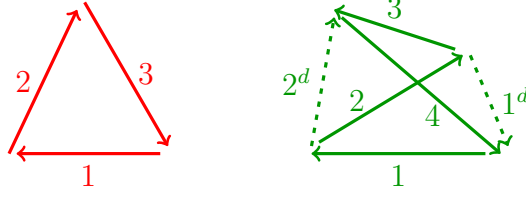


Figure 2.6: Diagrams corresponding to the 3D polyspectrum at **third** and **fourth** orders.

2.4.3 Plane wave expansion

The plane wave expansion, also called Rayleigh expansion, decomposes a 3D Fourier mode into the spherical and radial harmonic basis (see e.g. Brown 2009) :

$$\begin{aligned} e^{i\mathbf{k}\cdot\mathbf{x}} &= (4\pi) \sum_{\ell m} i^\ell Y_{\ell m}^*(\hat{k}) j_\ell(kr) Y_{\ell m}(\hat{n}) \\ &= (4\pi) \sum_{\ell m} i^\ell Y_{\ell m}(\hat{k}) j_\ell(kr) Y_{\ell m}^*(\hat{n}) \end{aligned} \quad (2.103)$$

where 3D vectors are decomposed into their radial and tangential parts : $\mathbf{k} = k\hat{k}$ and $\mathbf{x} = r\hat{n}$, and j_ℓ is the ℓ -order spherical Bessel function (see e.g. Brown 2009).

This equation allows us to decompose a 3D field and its correlation functions, into its radial and tangential components.

2.4.4 Projection of the past light cone

A frequent situation in cosmology is that the observation in a direction \hat{n} is the cumulative emission along the line-of-sight. Because the signal travels at a finite speed, the contribution at a distance r was emitted at a past time $t(r)$ ¹⁵. The space-time locus of these emissions is called the past light-cone of the observer. Hence an emission integrated along the past light-cone yields :

$$T(\hat{n}) = \int_0^\infty dr \frac{dT}{dr}(r\hat{n}, t(r)) \quad (2.104)$$

Noting for simplicity $f(\mathbf{x}, t) \equiv \frac{dT}{dr}$, decomposing f in Fourier, $T(\hat{n})$ in spherical harmonics and using the Rayleigh expansion, we find :

$$a_{\ell m} = i^\ell \int dr \frac{d^3\mathbf{k}}{2\pi^2} f(\mathbf{k}, t(r)) Y_{\ell m}^*(\hat{k}) j_\ell(kr) \quad (2.105)$$

Based on this equation, I have derived in Lacasa et al. (2013) the relation between the observed 2D polyspectrum and the emission 3D polyspectrum, if the latter is diagonal-independent. Indeed, taking the n -point correlation function :

$$\begin{aligned} \langle a_{\ell_1 m_1} \cdots a_{\ell_n m_n} \rangle_c &= i^{\ell_1 + \cdots + \ell_n} \int dr_{1\dots n} \frac{d^3\mathbf{k}_{1\dots n}}{2\pi^2} \left[Y_i^*(\hat{k}_i) j_{\ell_i}(k_i r_i) \right]_{i=1\dots n} \\ &\quad \times \mathcal{P}^{(n)}(\mathbf{k}_{1\dots n}, t_{1\dots n}) (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \cdots + \mathbf{k}_n) \end{aligned} \quad (2.106)$$

¹⁵Because of the expansion history of the universe the time $t(r)$ is not simply r/c . We have in fact $r(t) = \int_t^{t_0} \frac{cdt}{a(t)}$ where t_0 is now $a(t)$ is the scale factor and $a(t_0) = 1$ by convention.

Involving the cross-bispectrum between different epochs $t_1 \cdots t_n$ (with $t_i = t(r_i)$). If the polyspectrum is diagonal-independent we have :

$$\begin{aligned} \langle a_{\ell_1 m_1} \cdots a_{\ell_n m_n} \rangle_c &= i^{\ell_1 + \cdots + \ell_n} \int \left[dr_i \frac{k_i^2 dk_i}{2\pi^2} j_{\ell_i}(k_i r_i) \right]_{i=1 \cdots n} \mathcal{P}^{(n)}(k_{1 \cdots n}, t_{1 \cdots n}) \\ &\times \underbrace{\int d^2 \hat{\mathbf{k}}_{1 \cdots n} \left[Y_i^*(\hat{\mathbf{k}}_i) \right]_{i=1 \cdots n} (2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \cdots + \mathbf{k}_n)}_{\equiv A^{(n)}(\ell_{1 \cdots n}, m_{1 \cdots n}, \mathbf{k}_{1 \cdots n})} \quad (2.107) \end{aligned}$$

where, after calculations detailed in Lacasa et al. (2013), I find :

$$A^{(n)}(\ell_{1 \cdots n}, m_{1 \cdots n}, \mathbf{k}_{1 \cdots n}) = (4\pi)^n i^{\ell_1 + \cdots + \ell_n} \int x^2 dx [j_{\ell_i}(k_i x)]_{i=1 \cdots n} \times \underbrace{\int d^2 \mathbf{n} [Y_{\ell_i m_i}^*(\mathbf{n})]_{i=1 \cdots n}}_{\equiv \mathcal{G}(\mathbf{1} \cdots \mathbf{n})} \quad (2.108)$$

Hence, we find $\langle a_{1 \cdots n} \rangle_c \propto \mathcal{G}(\mathbf{1} \cdots \mathbf{n})$, which means that the 2D polyspectrum is also diagonal independent and we have :

$$\mathcal{P}^{(n)}(\ell_{1 \cdots n}) = (-1)^{\ell_1 + \cdots + \ell_n} \left(\frac{2}{\pi} \right)^n \int x^2 dx [dr_i k_i^2 dk_i j_{\ell_i}(k_i r_i) j_{\ell_i}(k_i x)]_{i=1 \cdots n} \mathcal{P}^{(n)}(k_{1 \cdots n}, t_{1 \cdots n}) \quad (2.109)$$

This expression can be further simplified using the Limber approximation. Indeed, if we assume that, as a function of k_i , the 3D polyspectrum varies slowly compared to the quickly oscillating Bessel functions, it can be factorised out of the k_i integral. Then, we can use the orthogonality relation of spherical Bessel functions, so that :

$$\begin{aligned} \int [k_i^2 dk_i j_{\ell_i}(k_i r_i) j_{\ell_i}(k_i x)]_i \mathcal{P}^{(n)}(k_{1 \cdots n}, t_{1 \cdots n}) &\simeq \mathcal{P}^{(n)}(k_{1 \cdots n}^*, t_{1 \cdots n}) \int [k_i^2 dk_i j_{\ell_i}(k_i r_i) j_{\ell_i}(k_i x)]_i \\ &\simeq \mathcal{P}^{(n)}(k_{1 \cdots n}^*, t_{1 \cdots n}) \left[\frac{\pi}{2x^2} \delta(r_i - x) \right]_i \end{aligned}$$

with the 3D polyspectrum evaluated at the peak of the Bessel function : $k_i^* = \frac{\ell_i + 1/2}{r_i}$. Hence, the r_i integrals are computed trivially and the 2D polyspectrum simplifies greatly :

$$\mathcal{P}^{(n)}(\ell_{1 \cdots n}) = (-1)^{\ell_1 + \cdots + \ell_n} \int \frac{x^2 dx}{x^{2n}} \mathcal{P}^{(n)}(k_{1 \cdots n}^*, t(x)) \quad (2.110)$$

which only involves the polyspectrum at a single epoch $t(x)$. Because of parity-invariance $\ell_1 + \cdots + \ell_n$ is even so that $(-1)^{\ell_1 + \cdots + \ell_n} = 1$.

Thus, the Limber approximation yields that the observed polyspectrum is an integral of the 3D polyspectrum over spherical shells. The weight x^{-2n} of these shells increases strongly as the radius decreases.

Remarks : Let us note that the Limber approximation is exact if the polyspectrum is constant, e.g., for a white noise.

For non-constant polyspectra, the Limber approximation may be expected to fail at low multipoles and most importantly at small radii x , as the spherical Bessel function are varying more slowly in these cases. The limit $x \rightarrow 0$ is particularly problematic as the weight x^{-2n} blows up while the 3D polyspectrum is supposed to go to zero as $k_i \rightarrow \infty$. Hence, this limit is prone to numerical instabilities, in particular at high orders n .

In this chapter, I have shown that cosmological analysis, either theoretical or empirical, is often more naturally set in harmonic space. Indeed, the correlation function introduced in the first chapter have a simpler form in harmonic space where they are called polyspectra. I have shown how these polyspectra may be estimated at second and third order, dealing with the problem of partial sky coverage, and how error bars on these estimations may be computed. I have thus described a major part of the framework necessary to characterise statistically celestial observations. This characterisation would then need to be compared to theoretical predictions. In a first step towards this goal, I have described in this chapter the projection of 3D polyspectra onto the sphere. Indeed, theoretical models are most naturally set in 3D space, as will be described in the two following chapters.

Chapter 3

Some cosmology

Contents

3.1	Friedman-Lemaître space-time	46
3.2	Inflation	47
3.2.1	General paradigm	47
3.2.2	Single field inflation	48
3.2.2.1	Background evolution	48
3.2.2.2	Perturbation generation and primordial non-Gaussianity	49
3.3	Cosmic Microwave Background anisotropies	50
3.4	Large Scale Structure	54
3.4.1	Linear growth	54
3.4.2	Local bias scheme	56
3.4.3	Halo model	58
3.4.4	A diagrammatic method for galaxy correlation functions	61

This chapter introduces the physical background motivating the application to cosmology of the statistical tools introduced in the previous chapters. I first describe the current paradigm for the generation of the primeval cosmological perturbations and why it may yield non-Gaussian signatures. I then introduce the main observables of present cosmology, which stem from these primeval perturbations : the Cosmic Microwave Background and the Large Scale Structure. For both of them, I sketch what non-Gaussianity studies may bring to their understanding and to the understanding of the primordial perturbation generation.

3.1 Friedman-Lemaître space-time

Observations of galaxy surveys and of the Cosmic Microwave Background show that the universe is isotropic (at least on very large scales). Combined with the Copernician principle stating that we do not live in a particular place in the universe, this leads to the conclusion that the universe is homogeneous (Peebles 1972; Peter & Uzan 2005).

These statements are meant to be statistical, that is *on average* the universe is homogeneous and isotropic. Of course this is not true in the exact realisation we live in, as e.g. a human being represents an overdensity $\delta\rho/\rho \sim 10^{30}$ compared to the average universe density. This average, through an ergodic hypothesis, can be thought as the mean over a sufficiently large space area. That is, if we smooth the density field over larger and larger windows, the field becomes asymptotically homogeneous.

Although homogeneous, we know since Hubble that the universe is not static. This led Alexander Friedman and independently Georges Lemaître to write down the most general metric (called FL metric in the following) describing a homogenous space-time :

$$ds^2 = -c^2 dt^2 + a^2(t) d\chi^2 \quad (3.1)$$

where $a(t)$ is the so-called scale factor which governs the distance between static observers. We normalise it to 1 at present time $a(t_0) = 1$. $d\chi^2$ is the comoving distance on spatial hypersurfaces. As today's observations constrain the global spatial curvature to be negligible (Planck Collaboration XVI 2013), we assume a euclidean metric $d\chi^2 = dx^2 + dy^2 + dz^2$.

From now on, unless explicitly specified, we work in natural units such that $c = 1$.

The cosmic time is noted t , but we may introduce another time coordinate η called conformal time, through $d\eta = \frac{dt}{a(t)}$. Photons travel along null geodesics : $0 = ds^2 = a^2(\eta)(-d\eta^2 + d\chi^2)$. Hence a photon travels a comoving distance $\chi = \eta$. This leads to the definition of the particle horizon (Rindler 1956) as the maximum distance that a photon may have travelled since the Big-Bang (defined at the end of this section) :

$$d_p(t) = a(t) \int_0^t \frac{dt'}{a(t')} \quad (3.2)$$

Observationally, the time of emission is most often inferred from the redshift $1+z = a(t_0)/a(t)$ which is an observable. In the following, a dot denotes derivative with respect to cosmic time (e.g. \dot{a}), while the occasional prime denotes derivative with respect to conformal time (e.g. a').

The statement of homogeneity and isotropy requires that the energy-impulsion tensor describing the matter content of the universe has the form $T^\mu_\nu = \text{diag}(-\rho, P, P, P)$, where ρ is the energy density and P the pressure of the matter content. The Einstein equations of General Relativity (GR) then govern the dynamics through two coupled differential equations called Friedman equations :

$$H^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} \quad (3.3)$$

$$\dot{\rho} + 3H(\rho + P) = 0 \quad (3.4)$$

with $H = \dot{a}/a$ being the Hubble parameter, G the gravitational constant and Λ the cosmological constant.

The differential system is closed by specifying the equation of state linking P and ρ . The matter content is typically composed of different fluids with equations of state $P = w\rho$, with $w = 0$ for baryonic and dark matter, $1/3$ for radiation, and the cosmological constant can be considered as a fluid with $w = -1$. If we index the different species by a label i , the first Friedman equation can be rewritten as an energy budget :

$$\sum_i \Omega_i = 1 \quad \text{where} \quad \Omega_i = \frac{\rho_i}{\rho_c} \quad (3.5)$$

with the critical density $\rho_c(z) = \frac{3H^2}{8\pi G}$.

Backward extrapolation of the Friedman equations with a standard matter content leads to an initial singularity : $a(t)$ hits zero in a finite cosmic time. This initial time, called the Big-Bang, is an essential singularity as geodesics cannot be continued through it. The current paradigm however is that GR ceases to be valid at early times when the temperature was of the order of the Planck temperature $T_P = \sqrt{\frac{\hbar c^5}{G k_B^2}}$ (with \hbar the reduced Planck's constant and k_B the Boltzmann constant). A quantum gravity theory (steming e.g. from Loop Quantum Gravity, string theory or non-commutative geometry) should in principle resolve the singularity. Thus the Big-Bang is an illegitimate extrapolation, which nevertheless serves as a referential origin.

3.2 Inflation

3.2.1 General paradigm

The Big-Bang model with standard energetic content (matter or radiation) yields a decelerated expansion. On the contrary, inflation is a proposed period of accelerated expansion in the very early universe, typically ending at an energy of the order 10^{16} GeV (Grand Unification scale). During this period the geometry is close to de Sitter meaning that the expansion is exponential $a(t) \propto e^{Ht}$. The introduction of inflation in the 80s was meant to solve several conceptual problems that arose with the standard Big-Bang model (Guth 1981; Hawking 1982) :

- flatness problem : the spatial curvature is observationally constrained to be negligible ($\Omega_k \ll 1$) today (Planck Collaboration XVI 2013). For a standard energetic content Ω_k grows with time, hence backward extrapolation would yield that spatial curvature becomes vanishingly small as time rolls back. This necessitates unnatural

initial conditions (fine-tuning problem). Inflation solves this problem by washing out spatial curvature through accelerated expansion, so that Ω_k becomes arbitrarily small at the end of inflation if that period lasted long enough.

- causal horizon : for a standard energetic content, the particle horizon at recombination (the time when the Cosmic microwave Background were last scattered, see Sect.3.3) is small compared to the typical distance between two points of the last scattering surface. This means that these points did not have time to communicate, let alone relax to the same thermal equilibrium. Nevertheless, these points have a thermal equilibrium emission with extremely close temperature (to 10^{-5} precision, see e.g. Strukov et al. 1992; Smoot et al. 1992). This again necessitates fine-tuned initial conditions which are unnatural. Inflation solves this problem, as the accelerated expansion means that these points were in fact much closer in a distant past, hence they had time to connect causally.
- monopole problem : most extensions of the standard model of particle physics predict the creation in the early universe of magnetic monopoles (an elementary particle with the equivalent of an electric charge but sourcing the magnetic field \vec{B} instead of \vec{E}). However no magnetic monopole has been observed to date in the universe. Inflation solves this apparent paradox by completely diluting the monopole density, so that there may be an arbitrarily small number of monopoles in the observable universe.

It has been realised rapidly that inflation also yields a process generating the perturbations of inhomogeneity (Mukhanov & Chibisov 1981; Linde 1982; Guth & Pi 1982; Starobinsky 1982). These are the very inhomogeneities that would later grow to form the CMB anisotropies, and much later on the large scale structure of the universe. In the next section, I illustrate briefly the implementation of inflation, and the generation of perturbations for the most standard inflationary model.

3.2.2 Single field inflation

The simplest implementation of inflation is mediated by a single scalar field $\varphi(\mathbf{x}, t)$ minimally coupled to general relativity. Thus, the action takes the form :

$$S = \frac{1}{16\pi G} \int \sqrt{-g} d^4x \left[R - \frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi - V(\varphi) \right] \quad (3.6)$$

where R is the Ricci scalar associated to the metric $g_{\mu\nu}$ and V is the potential of the scalar field.

3.2.2.1 Background evolution

We first assume homogeneity to derive the background equations, so that $\varphi(\mathbf{x}, t) = \varphi(t)$ and the metric is of the FL form.

The least-action principle gives the equations of motion. Namely by varying with respect to the metric $g_{\mu\nu}$ we get the Friedman equations, where the energy density and the

pressure of the scalar field are :

$$\rho(\varphi) = \frac{1}{2}\dot{\varphi}^2 + V(\varphi) \quad (3.7)$$

$$P(\varphi) = \frac{1}{2}\dot{\varphi}^2 - V(\varphi) \quad (3.8)$$

By varying the action with respect to φ , we get the Klein-Gordon equation :

$$\ddot{\varphi} + 3H\dot{\varphi} + \frac{dV}{d\varphi} = 0 \quad (3.9)$$

It is not the subject of this thesis to expose all the details of the inflationary dynamics. Let us just state that if the potential $V(\varphi)$ is sufficiently flat, the dynamics naturally enters a so-called slow-roll regime where the kinetic energy is small compared to $V(\varphi)$. In this slow-roll regime, the inflaton equation of state is $P \approx -\rho$. Thus, through Friedman equations, the expansion is quasi-exponential and hence satisfies the requirements of Sect.3.2.1. The deviations to this exponential behaviour (where the Hubble parameter is constant and the potential is flat) are parametrised by the so-called slow-roll parameters :

$$\epsilon = -\frac{\dot{H}}{H^2} \quad \text{and} \quad \delta = \frac{1}{3H^2} \frac{d^2V}{d\varphi^2} \quad (3.10)$$

We can note that these parameters also govern the magnitude of non-linearity in the Friedman and Klein-Gordon equations.

3.2.2.2 Perturbation generation and primordial non-Gaussianity

Inflation also generates the primordial density fluctuations around homogeneity. The basic idea is that quantum fluctuations of the inflaton are stretched out by the fast expansion and are frozen when their wavelength becomes larger than the Hubble length c/H . The details are complex and would necessitate a long description (see e.g. Peter & Uzan 2005). First the FL metric must be perturbed with gauge-invariant quantities. Then the action Eq.3.6 must be quantized around the background, and quantum initial conditions must be chosen (the standard choice being the Bunch-Davies vacuum, which is Gaussian). Nevertheless, one can show that the resulting density perturbations $\delta\rho/\rho$ (related to the gravitational potential Φ through the Poisson equation) have a nearly scale-invariant power spectrum :

$$P_\delta(k) \propto k^{n_S} \quad \text{with} \quad n_S \approx 1 \quad (3.11)$$

More precisely it would be exactly scale invariant ($n_S = 1$, so that the dimensionless gravitational power spectrum $\frac{k^3 P_\Phi(k)}{2\pi^2} \propto k^{-1} P_\delta(k)$ is constant) if the background evolution was de Sitter, i.e. $a(t) \propto e^{Ht}$. The deviation to scale invariance is linked to this deviation from de Sitter space, i.e. to the slow-roll parameters :

$$n_S - 1 = 2\delta - 4\epsilon \quad (3.12)$$

This is a general prediction of inflation models (Mukhanov & Chibisov 1981; Hawking 1982; Guth & Pi 1982), which yield a spectral index close to, but smaller than, unity ; and

indeed Planck has revealed a near scale invariant spectrum, detecting significantly the deviation to scale invariance (at 68% C.L Planck CMB alone gives $n_s = 0.9616 \pm 0.0094$ (Planck Collaboration XVI 2013)).

As for the background, the slow roll parameters govern the magnitude of non-linearity in the perturbation evolution equations; as the vacuum initial condition is Gaussian, it is not surprising that the resulting density perturbation are Gaussian to leading order and that non-Gaussianity scales proportionally to the slow-roll parameters. Indeed, Maldacena (2003) has shown that the bispectrum amplitude in the squeezed limit ($k_1 \ll k_2 \approx k_3$) is :

$$B_\Phi(k_1, k_2, k_3) = \frac{5}{3} \frac{(1 - n_s)}{2} [P_\Phi(k_1)P_\Phi(k_2) + P_\Phi(k_1)P_\Phi(k_3) + P_\Phi(k_2)P_\Phi(k_3)] \quad (3.13)$$

where the factor $\frac{5}{3}$ comes from the relation between the gravitational potential and the curvature perturbation.

Later on, Creminelli & Zaldarriaga (2004) have shown in fact that this equation holds in the squeezed limit for a whole class of inflation models, e.g., beyond slow-roll, as long as the inflaton is the only dynamical field (single clock). Hence, detecting primordial non-Gaussianity in the squeezed limit would rule out whole classes of inflationary models (*ibid.*; Komatsu 2002).

In a general way, for inflation to generate non-Gaussian primordial perturbations (where the non-Gaussianity may peak in configurations other than squeezed, e.g. in equilateral), we need either the initial conditions to be non-Gaussian (non Bunch-Davies vacuum) or the evolution to be non-linear. The latter possibility is the case for exotic models with e.g., non-canonical kinetic term, multi-field, varying speed of sound etc (see e.g. Bartolo et al. 2004; Renaux-Petel 2009; Lewis 2011, and references therein). This further motivates the search for primordial non-Gaussianity as it may discriminate inflation models. Indeed these models are degenerate at the power spectrum level, as they are all built to reproduce near scale invariance.

3.3 Cosmic Microwave Background anisotropies

The Cosmic Microwave Background (CMB), predicted by Gamow (1948); Dicke et al. (1965) and then discovered by Penzias & Wilson (1965), is the relic radiation of the primordial plasma. Namely in the early universe (between $z \sim 10^8$ and 10^3), baryons and photons were tightly coupled into a plasma through Compton scattering of the numerous photons ($\sim 10^9$ per baryon) off free electrons. Due to the high density of free electrons, the photon mean free path was small and the universe was opaque. Baryonic matter and radiation were in local thermal equilibrium and experienced acoustic oscillations through the competition of gravitational infall and radiation pressure, while dark matter was thermally decoupled as it interacts only gravitationally.

With expansion the universe cooled down, at a redshift $z_{\text{rec}} \sim 1000$ and a temperature $T \simeq 3000$ K, the photons were no more energetic enough to prevent the binding of electrons with available nuclei (in particular hydrogen and helium). This event is called recombination and leads to a dramatic decrease of the ionisation fraction so that photons could free stream in the universe (or nearly so, until reionisation took place and scattered some of them). These numerous photons still represent the vast majority of radiation

energy in the universe today (see e.g. Fig.4.1), and make up the CMB. The expansion has stretched their wavelength to the microwave band today, and as this stretching is achromatic they have kept a blackbody spectral distribution albeit with a current temperature of 2.725 K (Fixsen & Mather 2002).

The perturbations created by inflation lead to density, velocity and gravitational potential inhomogeneities in the baryon-photon plasma. These tiny inhomogeneities in the plasma at recombination lead to tiny temperature anisotropies (with relative amplitude $\sim 10^{-5}$) in the observed CMB today. There are several competing physical effects leading to the temperature anisotropies, and this can be summarised into an equation linking the initial gravitational potential fluctuations (where initial means end of inflation) to the harmonic coefficients of the observed relative temperature fluctuations :

$$a_{\ell m}(\eta_0) = \int \frac{d^3\mathbf{k}}{2\pi^2} i^\ell \Delta_\ell(k, \eta_i, \eta_0) \Phi(\mathbf{k}, \eta_i) Y_{\ell m}(\hat{k}) \quad (3.14)$$

where η_0 is today's conformal time, $\Phi(\mathbf{k}, \eta_i)$ is the Fourier transform of the gravitational potential at the initial time (after inflation), and $\Delta_\ell(k)$ is the transfer function which encodes all the physics and can be computed by a Boltzmann code, e.g. CAMB (Lewis & Bridle 2002) or CLASS (Blas et al. 2011).

For example, on large scales the Sachs-Wolfe effect (Sachs & Wolfe 1967) is dominant. In the instantaneous recombination approximation it yields $\frac{\Delta T}{T}(\hat{n}) = -\frac{\Phi(r_{\text{rec}} \hat{n})}{3}$, and the transfer function takes the form :

$$\Delta_\ell(k) \simeq -\frac{1}{3} j_\ell(kr_{\text{rec}}) \quad (3.15)$$

up to growth factors (explained in Sect.3.4.1) giving the evolution of potentials between η_i and η_{rec} for frozen super-Hubble modes.

From Eq.3.14, we can then derive the CMB power spectrum :

$$C_\ell(\eta_0) = \frac{2}{\pi} \int k^2 dk \Delta_\ell^2(k, \eta_i, \eta_0) P_\Phi(k) \quad (3.16)$$

Figure 3.1 shows the CMB power spectrum measured by *Planck* (Planck Collaboration XV 2013). Specifically they plotted $\mathcal{D}_\ell = \frac{\ell(\ell+1)}{2\pi} C_\ell$, a now standard convention. Indeed for scale invariant fluctuations the Sachs-Wolfe effect produces $C_\ell \propto \frac{1}{\ell(\ell+1)}$ at low multipoles, and this convention reduces to the power per logarithmic multipoles bin $\frac{\ell(2\ell+1)}{4\pi} C_\ell$ when $\ell \gg 1$.

The power spectrum exhibits distinct characteristics which are the imprints of the primordial physics :

- the Sachs-Wolfe plateau at low multipoles previously explained.
- so-called acoustic peaks, which correspond to a dip in the real-space 2-point c.f. at an angle $\theta_s \simeq 1.2^\circ$. θ_s is the angle subtended by the sound horizon at last scattering, i.e. the distance sound waves traveled in the photon-baryon plasma between the generation of density perturbations and recombination.
- so-called Silk damping at high multipoles (Silk 1968), as the diffusion of photons during the period of recombination erases their inhomogeneity.

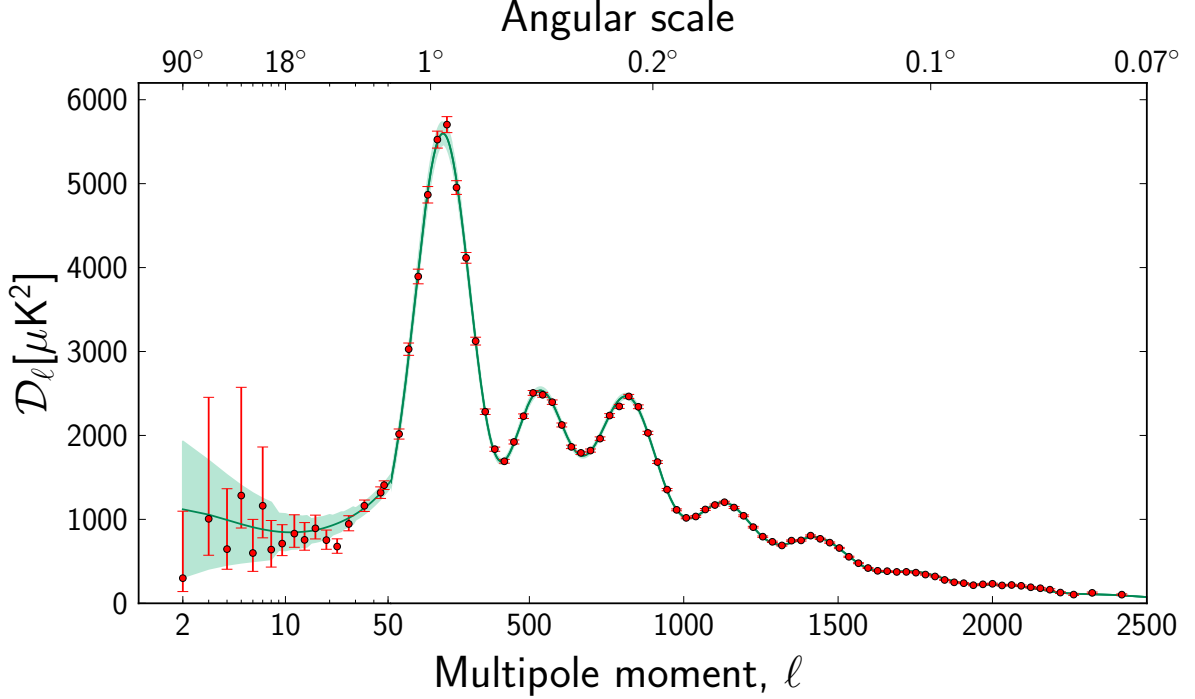


Figure 3.1: *Planck* CMB power spectrum (red dots, binned) along with the best fit cosmological model and its cosmic variance (Planck Collaboration XV 2013). Note the log-linear x-axis, and the y-axis is $\mathcal{D}_\ell = \frac{\ell(\ell+1)}{2\pi}C_\ell$.

The transfer function $\Delta_\ell(k)$ encodes all these effects; as such it is highly sensitive to the primordial physics, hence to the cosmological parameters. The CMB is thus a powerful probe of cosmological parameters, and has been observed by several experiments in this prospect (e.g. Sievers et al. 2003; MacTavish et al. 2006; Komatsu et al. 2011); the *Planck* results have put the most stringent constraints to date on them (Planck Collaboration XVI 2013).

Higher order CMB polyspectra may also be interesting, to hunt for primordial non-Gaussianity. For example, from Eq.3.14 and assuming a primordial non-Gaussianity shape, one can derive the CMB bispectrum. For data analysis, the most popular non-Gaussianity shape is the local form, parametrised by a non-linearity parameter f_{NL} :

$$\Phi(\mathbf{x}) = \Phi_G(\mathbf{x}) + f_{\text{NL}} (\Phi_G(\mathbf{x})^2 - \langle \Phi_G(\mathbf{x})^2 \rangle) \quad (3.17)$$

where Φ_G is a Gaussian field. To the first order in f_{NL} , this creates a gravitational bispectrum of the form :

$$B_\Phi(k_1, k_2, k_3) = f_{\text{NL}} [P_\Phi(k_1)P_\Phi(k_2) + 5 \text{ perm.}] \quad (3.18)$$

i.e. the same form as Eq.3.13¹ so that for standard inflation $f_{\text{NL}} = \frac{5}{12}(1 - n_s)$ (greater value of f_{NL} may be generated e.g. by multifield models (Bartolo et al. 2004)).

¹Although Eq.3.13 was only valid in the squeezed limit, the bispectrum peaks in these configurations so that this form captures the dominant behaviour.

With calculations similar to that of Sect.2.4.4, one can show that the CMB bispectrum takes the form :

$$b_{\ell_{123}} = f_{\text{NL}} \int r^2 dr \alpha_{\ell_1}(r) \beta_{\ell_2}(r) \beta_{\ell_3}(r) + 5 \text{ perm.} \quad (3.19)$$

with :

$$\alpha_{\ell}(r) = \frac{2}{\pi} \int k^2 dk \Delta_{\ell}(k) j_{\ell}(kr) \quad (3.20)$$

$$\beta_{\ell}(r) = \frac{2}{\pi} \int k^2 dk P_{\Phi}(k) \Delta_{\ell}(k) j_{\ell}(kr) \quad (3.21)$$

r can be thought of as the comoving distance on the past light cone (although in the derivation it is just a bound variable), and indeed the filters α_{ℓ} and β_{ℓ} peak at recombination $r \simeq r_{\text{rec}}$. $\frac{\beta_{\ell}(r)}{C_{\ell}}$ is the optimal linear filter reconstructing the gravitational potential at comoving distance r from the observed CMB $a_{\ell m}$ (Komatsu et al. 2005).

Other forms of primordial non-Gaussianity, as long as they produce a gravitational bispectrum of the form $B(k_{123}) = \sum_i f_i(k_1) g_i(k_2) h_i(k_3) + \text{perm.}$, yield a CMB bispectrum with a functional form similar to Eq.3.19 albeit with different filters. For example, two other popular forms of non-Gaussianity are the so-called equilateral and orthogonal ones, which respectively arise for models with non-canonical kinetic terms or with non-standard vacuum initial conditions (see e.g., Bartolo et al. 2004; Renaux-Petel 2009; Lewis 2011). The CMB bispectrum that they yield involves two new filters traditionally noted $\gamma_{\ell}(r)$ and $\delta_{\ell}(r)$, see e.g. Curto et al. (2011) for details.

Non-Gaussianity in the observed CMB is also generated by non-linear evolution (Pitrou et al. 2010; Huang & Vernizzi 2012) and by the secondary anisotropies, that is, anisotropies due to the distortion of the primordial CMB by the large scale structure along the line of sight. These secondary anisotropies will be reviewed in Sect.4.1. In particular, the iSW-lensing-lensing cross-bispectrum (shortened to ISW-lensing hereafter) has been shown to be important to consider for CMB NG studies (e.g. Smith & Zaldarriaga 2011; Hanson et al. 2009), and has indeed been measured by *Planck* as will be explained in Sect.5.2.2.

Figure 3.2 shows the theoretical CMB bispectrum for local type primordial non-Gaussianity and $f_{\text{NL}} = 1$, plotted in the parametrisation proposed in Lacasa et al. (2012) and detailed in Sect.2.2.4.6). The transfer function was computed using CAMB (Lewis & Bridle 2002).

Note that, for visualisation purpose, the color code is highly non-linear to highlight all features of the bispectrum. Each subplot corresponds to a perimeter bin; in that subplot the upper left corner corresponds to squeezed triangles, upper right to equilateral and lower right to folded triangles. The perimeter value, $\ell_1 + \ell_2 + \ell_3$, is indicated in the lower left corner of each subplot.

The bispectrum is dominated by the squeezed configurations where it is negative. Indeed one of the multipoles is small so that the transfer function is negative (Sachs-Wolfe regime $\frac{\Delta T}{T} = -\frac{\Phi}{3}$) while the two other multipoles are near equal so that their transfer function has same sign. The complex pattern which develops at increasing multipoles is the consequence of the acoustic peaks which change the sign of the transfer function. These sign changes could not be seen in the power spectrum as it is intrinsically positive. They occur at multipoles corresponding to local minima of C_{ℓ} and can be seen easily in the equilateral configurations of the bispectrum. In absolute value, the bispectrum is of order

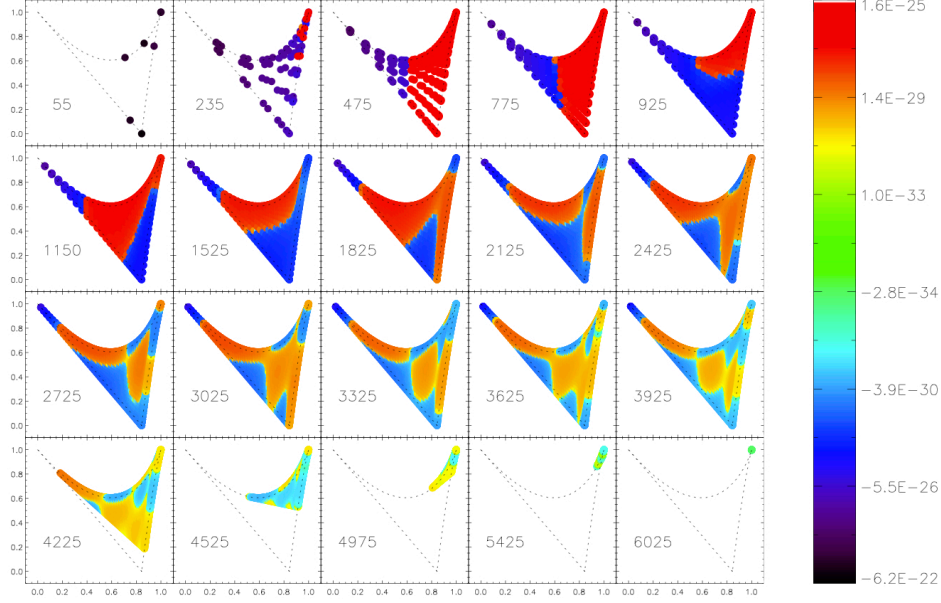


Figure 3.2: Theoretical CMB bispectrum for local type primordial non-Gaussianity and $f_{\text{NL}} = 1$. The perimeter value $\ell_1 + \ell_2 + \ell_3$ is indicated in each subplot. The color code is highly non-linear. The bispectrum peaks in the squeezed configurations with negative values (black-violet colors). Note that the maximum resolution condition $\ell \leq \ell_{\text{max}}$ restrains the possible triangle shape at high perimeters, e.g. only near equilateral triangles are present in the last bin $P \sim 3\ell_{\text{max}}$

$\mathcal{O}(10^{-20})$ in the squeezed triangles, as can be expected from Eq.3.17 or 3.18, given that the perturbations in Φ are of order $\mathcal{O}(10^{-5})$.

3.4 Large Scale Structure

The perturbations created by inflation and visible in the CMB grow through gravitational infall to form the structure of the universe that we see today (Peter & Uzan 2005; Bernardeau 2007). This section describes the growth of such structures and how to model the large scale structure of the universe.

3.4.1 Linear growth

A newtonian approach is valid on sub-Hubble scales ($\lambda < c/H$) where general relativity effects are negligible (see e.g. Weinberg (1972) for a much more complete discussion). If a single pressureless fluid dominates the inhomogeneities, e.g., in matter- or Λ -dominated era; the Poisson equation, mass conservation and Euler equation take the form :

$$\Delta_{\text{p}}\Phi = 4\pi G\rho \quad (3.22)$$

$$\left.\frac{\partial\rho}{\partial t}\right|_{\text{p}} + \text{div}_{\text{p}}(\rho\vec{v}) = 0 \quad (3.23)$$

$$\rho \left. \frac{\partial \vec{v}}{\partial t} \right|_p + \rho(\vec{v} \cdot \vec{\nabla}_p) \vec{v} = -\rho \vec{\nabla}_p \Phi \quad (3.24)$$

where \vec{v} is the velocity field, and the subscript p denotes that derivatives are taken with respect to physical positions $\mathbf{x} = a(t)\mathbf{r}$, where \mathbf{r} is the comoving position. The physical and comoving time-derivative are also different, as the time-derivative is meant to be taken at constant spatial coordinates. The conversion between physical and comoving derivatives, the latter being noted with a subscript c, are :

$$\vec{\nabla}_p = \frac{1}{a} \vec{\nabla}_c \quad (3.25)$$

$$\left. \frac{\partial}{\partial t} \right|_p = \left. \frac{\partial}{\partial t} \right|_c - H \mathbf{r} \cdot \vec{\nabla}_c \quad (3.26)$$

We will note $\delta_m = \frac{\rho - \bar{\rho}}{\bar{\rho}}$ the relative perturbation of matter density, comprising baryonic matter and dark matter, and $\delta \vec{v}$ the velocity perturbation. If we linearise Eq.3.22-3.23-3.24 around the background solution² ($\rho = \bar{\rho}(t)$, $\vec{v} = H\mathbf{x}$, $\Phi = 0$), we get in comoving coordinates :

$$\Delta_c \Phi = 4\pi G a^2 \bar{\rho} \delta_m \quad (3.27)$$

$$\left. \frac{\partial \delta_m}{\partial t} \right|_c + \frac{1}{a} \vec{\nabla}_c \cdot \delta \vec{v} = 0 \quad (3.28)$$

$$\left. \frac{\partial \delta \vec{v}}{\partial t} \right|_c + H \delta \vec{v} = -\frac{1}{a} \vec{\nabla}_c \Phi \quad (3.29)$$

These equations can be combined to obtain a second-order differential equation on δ_m only :

$$\ddot{\delta}_m + 2H \dot{\delta}_m = 4\pi G \bar{\rho} \delta_m \quad (3.30)$$

where, for simplicity, the comoving time-derivative is noted with an overdot. As there are only time derivatives involved, this equation has separable solutions, i.e. of the form $\delta_m(\mathbf{x}, t) = D(t)f(\mathbf{x})$. Moreover, the differential equation is of second-order, so there are two independent solutions, with amplitude fixed by the initial conditions :

$$\delta_m(\mathbf{x}, t) = D^+(t) \delta_m^+(\mathbf{x}, t_i) + D^-(t) \delta_m^-(\mathbf{x}, t_i) \quad (3.31)$$

Let us note that the right-hand side of Eq.3.30 can also be written $\frac{3}{2}H^2\delta_m$.

Using the Friedmann equations 3.3 & 3.4, we see that the Hubble parameter H is solution of Eq.3.30, hence $D^-(t) = H(t)/H(t_i)$. The other mode can be searched for in the form $D^+(t) = H(t)g(t)$, which leads to a first order differential equation on \dot{g} . Hence we find :

$$D^+(t) \propto H(t) \int_{t_i}^t \frac{dt}{[a(t)H(t)]^2} = H \int_{a_i}^a \frac{da}{(aH)^3} \quad (3.32)$$

During matter domination $D^+ \propto a \propto t^{2/3}$ and $D^- \propto t^{-1}$, so that the corresponding solutions are respectively called the growing mode and the decaying mode.

²Except that the background variables are not solutions of the system, as they do not follow Newtonian dynamics but GR. Nevertheless, the Newtonian approach to perturbations produces the same perturbations equations as GR.

As long as the matter perturbations are small, they grow linearly without mixing Fourier modes. No extra non-Gaussianity is generated and the correlation functions of the density field are kept unchanged ; this stands as well for the velocity and gravitational potential fields.

However, perturbations eventually grow out of the linear regime, and the non-linear dynamic is not solvable analytically. As the non-linear terms in Eq.3.22-3.23-3.24 all involve spatial derivatives, the non-linearity affects predominantly small scales. In the Fourier range, large scales are hence still in the linear regime while small scales are non-linear. As time goes by, the non-linear range extends to larger and larger scales.

The non-linearity can be tackled perturbatively, e.g. with Eulerian or Lagrangian perturbation theory (see Bernardeau et al. 2002, for a review). The linear approach described previously then gives the first order solution, and the system is solved iteratively order by order. Higher orders of computation, resummation of loops contributions etc, enable the domain of validity to extend to smaller scales. But perturbation theory eventually fails in the fully non-linear regime (Buchert 1993). The purpose of the next two sections is to describe the current approach used to tackle the distribution of matter both at large and small scales, and in particular the distribution of galaxies.

3.4.2 Local bias scheme

At late times, the large scale structure of the universe forms a complex web : at large scales filaments of dark matter are separated by large void areas, haloes form in these filaments and host clusters of galaxies, the largest ones forming at the filaments intersections. For illustration, Fig.3.3 shows a slice of a N-body dark matter simulation runned by the consortium DEUS (Rasera et al. 2010), and Fig.3.4 shows the galaxy distribution observed by the Sloan Digital Sky Survey (Tegmark et al. 2004).

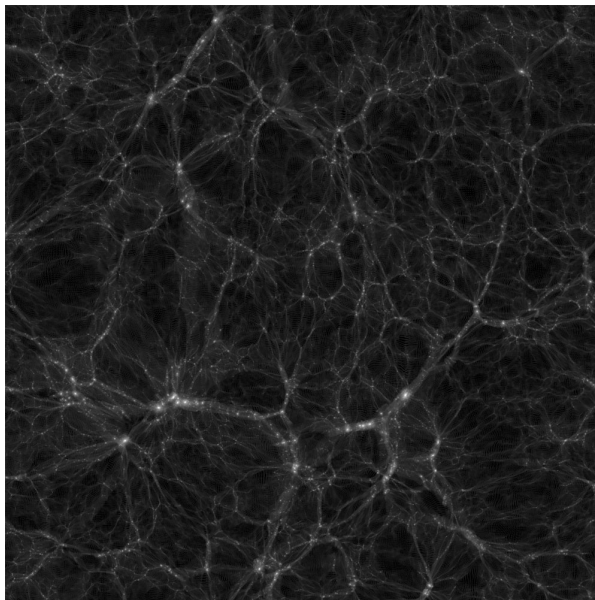


Figure 3.3: Slice at $z=0$ of a dark matter simulation runned by the DEUS consortium (Rasera et al. 2010), showing the distribution of dark matter in filaments, voids and haloes.

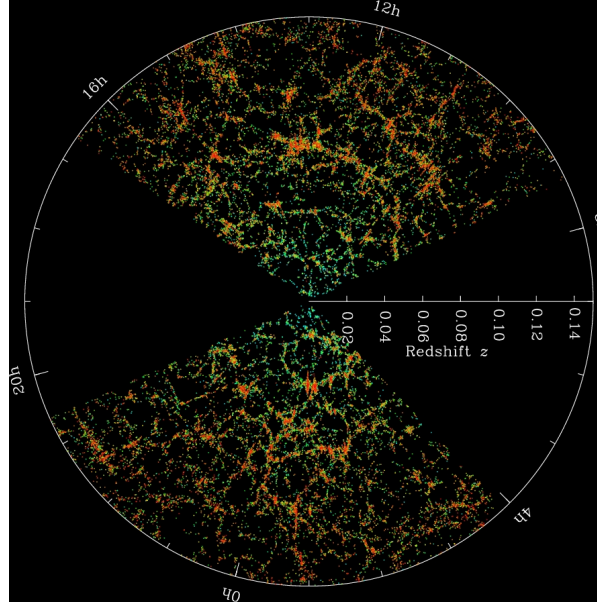


Figure 3.4: Galaxy distribution observed by the Sloan Digital Sky Survey (Tegmark et al. 2004).

The study of the galaxy distribution in the universe is a long-standing field in cosmology (e.g. Peebles 1972). Their clustering has been particularly studied from the point of view of the real-space 2-point correlation function (*ibid.*). Early observations showed that the galaxy correlation function was on large scale a biased version of the matter correlation function : $\zeta_g(r) = b^2 \zeta_m(r)$ ³. This lead Fry & Gaztanaga (1993) to propose that galaxies are a tracer of matter through a local non-linear transformation : $\delta_g(\mathbf{x}) = F(\delta_m(\mathbf{x}))$. This transformation may be Taylor-expanded to yield the so-called local biasing scheme :

$$\delta_g(\mathbf{x}) = \sum_{k=0}^{+\infty} \frac{b_k}{k!} \delta_m^k(\mathbf{x}) \quad (3.33)$$

where $b_0 = -\sum_{k=2}^{+\infty} \frac{b_k}{k!} \langle \delta_m^k \rangle$ so that $\langle \delta_g \rangle = 0$. b_i are the bias parameters, the most studied being b_1 also known as the linear bias. As the relation Eq.3.33 is local, the bias parameters can be inferred from data, simulations or theoretical predictions solely from the 1-point function, i.e., the abundances.

This functional form holds on large scales for any selected objects, e.g., red galaxies, groups etc, and allows to describe their large scale distribution. However on smaller scales, the main driver of galaxy clustering is whether or not they live in the same halo. As haloes are the locus of galaxy formation, they are the key ingredient to model the galaxy distribution. The number and properties of the hosted galaxies are a function of the physical parameters of the halo, and in particular of the halo mass. It has been demonstrated by Manera & Gaztañaga (2011) that, for haloes above a certain mass threshold, the halo distribution was well fitted by the local bias scheme with bias parameters derived from the halo abundances (also known as the halo mass function, described in the next section).

³This form is also expected for a thresholded Gaussian random field (Kaiser 1984).

Thus the density field of haloes δ_h at a given mass M follows :

$$\delta_h(\mathbf{x}|M) = \sum_{k=0}^{+\infty} \frac{b_k(M)}{k!} \delta_{DM}(\mathbf{x})^k \quad (3.34)$$

where δ_{DM} is the dark matter density field, the statistics of which are predicted through perturbation theory.

Note that this description breaks down on small scale when non-linearities become important.

3.4.3 Halo model

Press & Schechter (1974) pioneered the halo description of the Large scale Structure with a spherical collapse model. They gave an analytical formula for the abundance of haloes per mass bin and per comoving volume at a given redshift $\frac{dn_h}{dM dV_c}$, where they found the functional form :

$$\frac{dn_h}{dM dV_c}(M, z) \propto M^{-1-\alpha} \exp[-\text{cste}(1+z)^2 M^{2-2\alpha}] \quad (3.35)$$

where α is a parameter between $\frac{1}{2}$ and $\frac{1}{3}$ (Press & Schechter 1974).

This functional form highlights the universal properties of the mass function : a relatively flat slope at low masses and a sharp cut-off at high masses, here with an exponential decay. Hence massive haloes are rare while low-mass haloes are numerous. The cut-off decreases with increasing redshift, so that massive haloes are even rarer at early times. Indeed they did not have enough time to form through gravitational infall.

Let us note that the exponential decay in Eq.3.35 is the consequence of the assumption of Gaussian initial conditions, as massive haloes correspond to high extrema of the initial density field. Indeed, the exponential is written as $\exp\left[-\frac{\delta_c^2(z)}{2\sigma^2(M,z)}\right]^4$. Hence, non-Gaussian initial conditions would impact the mass function ; massive haloes are particularly sensitive to this possible effect and may be a probe of primordial NG (e.g., Sefusatti 2005).

Since Press & Schechter (1974), more realistic mass functions have been proposed, beyond the spherical collapse model, with up to date cosmological parameters, and in agreement with N-body simulations. For this work I have been using the mass function proposed by Sheth & Tormen (1999). Although it is not the most recent mass function, it is the only one for which the bias parameters beyond first order are available. In Fig.3.5 the Sheth & Tormen (1999) mass function and associated bias functions are displayed at $z=0$ and 2 over a range of masses representative of galaxy clusters.

As noted by Cooray & Sheth (2002), the average first order bias is one while higher order biases average to zero :

$$\int dM \frac{dn_h}{dM} \frac{M}{\bar{\rho}} b_1(M) = 1 \quad (3.36)$$

$$\int dM \frac{dn_h}{dM} \frac{M}{\bar{\rho}} b_n(M) = 0 \quad \forall n > 1 \quad (3.37)$$

⁴ δ_c is the critical density, such that a spherical overdensity with that contrast collapses through gravitational infall instead of expanding with the Hubble flow. $\sigma(M, z)$ is the variance of the density field smoothed over a length R with $M = \frac{4}{3}\pi\bar{\rho}R^3$

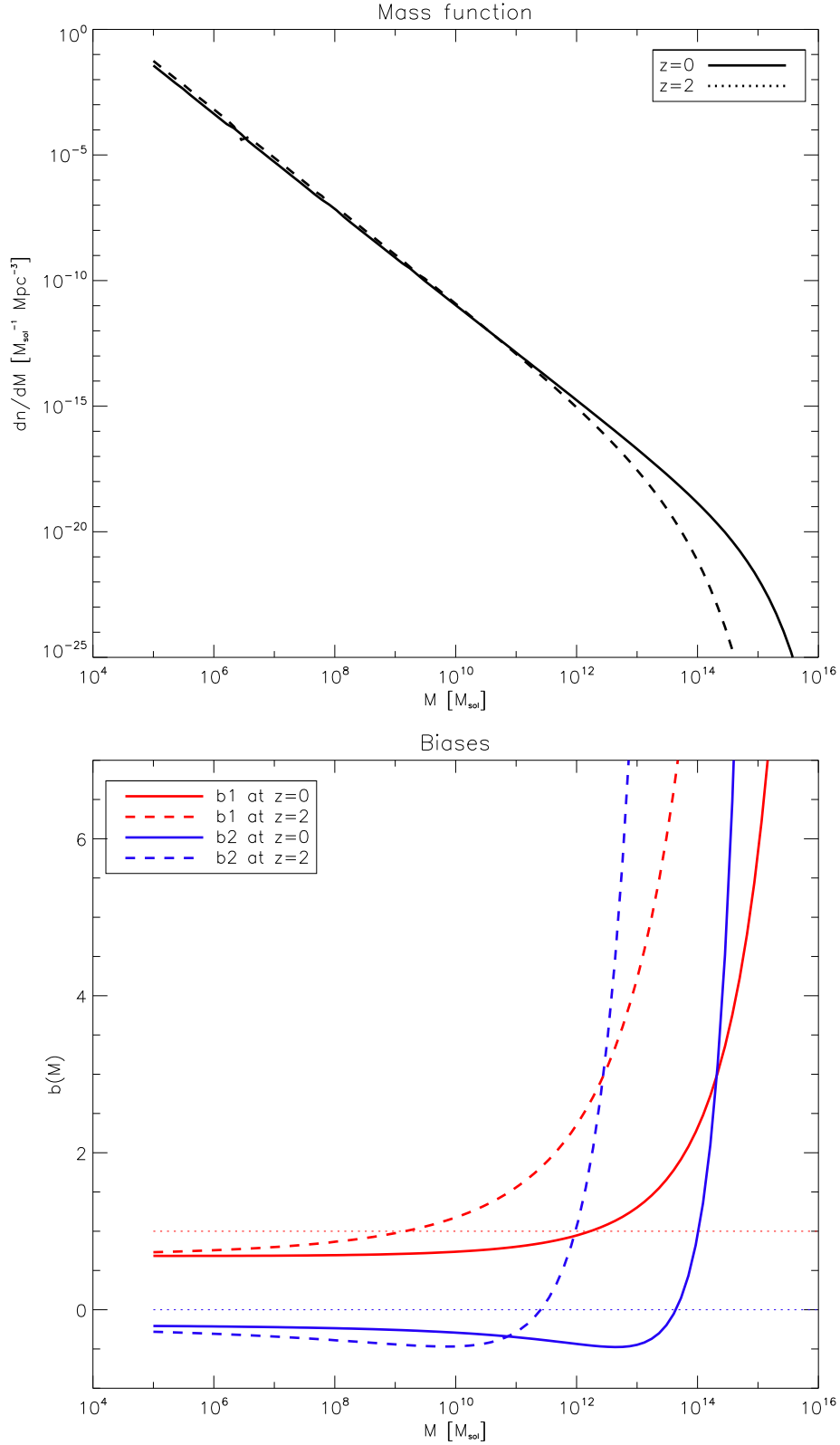


Figure 3.5: *Upper panel* : Sheth & Tormen (1999) mass function. *Lower panel* : associated bias parameters at **first** and **second** order, at $z=0$ (solid lines) and $z=2$ (dashed lines); dotted lines show the mass averaged biases, see Eq.3.36&3.37.

These averages are shown as dotted lines in Fig.3.5. We see that massive haloes are more biased than low-mass haloes. At higher redshifts, massive haloes are even more biased as they become extreme and very rare objects.

To describe the inner halo, we will assume it is spherical and is hence characterised by a radial profile. Numerical simulations show that dark matter haloes follow a universal profile named NFW for Navarro, Frenk and White (see Navarro et al. 1997) :

$$\rho(r|M) = \frac{\rho_s}{(r/r_s)(1 + r/r_s)^2} \quad (3.38)$$

where $\rho_s(M)$ and $r_s(M)$ (the scale radius) are the parameters of the model. These parameters are related to the virial radius r_{vir} ⁵ and virial mass M_{vir} of the halo, and in particular one defines the concentration parameter $c(M, z)$ such that $r_{\text{vir}} = cr_s$. In the following we adopt the mean concentration parameter given by Dolney et al. (2004), and neglect scatter in this parameter.

We will also assume that the galaxy distribution follows the normalised dark matter profile $u(r|M)$ ⁶. While this may be controversial at first sight, we expect it to be a good approximation as galaxy formation is driven by initial overdensities in the dark matter density field.

The last element needed to describe the galaxy distribution is the number of galaxies inside a given halo. This is called the Halo Occupation Distribution (HOD hereafter). Following observations and numerical simulations (Berlind & Weinberg 2002; Zheng et al. 2005; van den Bosch et al. 2007), the number of galaxies is typically described as the sum of a central galaxy and possibly satellite galaxies : $N_{\text{gal}} = N_{\text{cen}} + N_{\text{sat}}$ where both numbers are random variables drawn from a distribution depending on the halo mass. N_{cen} is either 0 or 1 with a mean value being a step-like function (Kravtsov et al. 2004). N_{sat} is drawn, conditionally to $N_{\text{cen}} = 1$, from a Poisson distribution (Zheng et al. 2005) whose mean is a power law with the halo mass at high masses. Following Tinker & Wetzel (2010); Pénin et al. (2012), we take :

$$\langle N_{\text{cen}} \rangle = \frac{1}{2} \left[1 + \text{erf} \left(\frac{\log M - \log M_{\text{min}}}{\sigma_{\log M}} \right) \right] \quad (3.39)$$

$$\langle N_{\text{sat}} \rangle = \frac{1}{2} \left[1 + \text{erf} \left(\frac{\log M - \log 2M_{\text{min}}}{\sigma_{\log M}} \right) \right] \left(\frac{M}{M_{\text{sat}}} \right)^{\alpha_{\text{sat}}} \quad (3.40)$$

where M_{min} , M_{sat} , $\sigma_{\log M}$ and α_{sat} are the four HOD parameters. M_{min} is the typical mass for which a halo starts to host a central galaxy, $\sigma_{\log M}$ governs the width of the step-like function for this central galaxy, M_{sat} is the typical mass above which a halo hosts a satellite galaxy, and α_{sat} governs the slope of the number of satellite galaxies at high masses. We take as fiducial values $M_{\text{min}} = 10^{11.5} M_{\odot}$, $M_{\text{sat}} = 10 M_{\text{min}}$, $\sigma_{\log M} = 0.65$ and $\alpha_{\text{sat}} = 1.4$ as found by Cosmic Infrared Background studies (Viero et al. 2009; Amblard et al. 2011; Planck-Collaboration 2011; Pénin et al. 2012). Heuristically, we

⁵The virial radius is the radius such that inner particles can be considered gravitationally bound together so that the virial theorem applies; we consider $r_{\text{vir}} = r_{200}$ the radius within which the average density is 200 times the critical density.

⁶ $u(r|M) = \frac{\rho(r|M)}{\int d^3\mathbf{r} \rho(\mathbf{r}|M)}$

would expect the number of galaxies to grow proportionally to the halo mass so that $\alpha_{\text{sat}} = 1$. However, CIB observations typically find best-fit values $\alpha_{\text{sat}} > 1$; this may be a modeling shortcoming. Figure 3.6 shows the resulting average galaxy number depending on the halo mass.

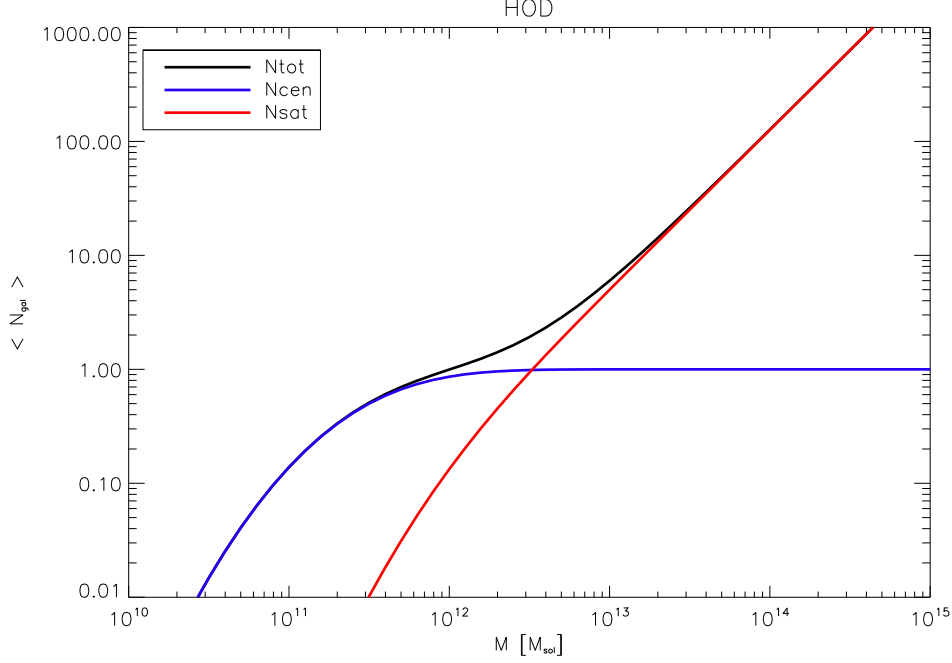


Figure 3.6: Average number of galaxies in a halo depending on its mass, with the HOD described in Eq.3.39&3.40 and fiducial HOD parameters $M_{\text{min}} = 10^{11.5} M_{\odot}$, $M_{\text{sat}} = 10 M_{\text{min}}$, $\sigma_{\log M} = 0.65$ and $\alpha_{\text{sat}} = 1.4$.

3.4.4 A diagrammatic method for galaxy correlation functions

This section assembles the physical ingredients of the halo model described previously and proposes a new diagrammatic formalism allowing to compute the galaxy correlation functions/polyspectra.

The redshift dependence of all functions will be implicit in this whole section to lighten notations. The galaxies number per comoving volume at a given point \mathbf{x} reads :

$$n_{\text{gal}}(\mathbf{x}) = \sum_i n_{\text{gal}}(\mathbf{x}|i) \quad (3.41)$$

with i being the halo index.

In the literature, $n_{\text{gal}}(\mathbf{x}|i)$ is assumed, implicitly or explicitly, to be a smooth distribution following the halo profile :

$$n_{\text{gal}}(\mathbf{x}|i) = N_{\text{gal}} u(\mathbf{x}|M_i) \quad (3.42)$$

However galaxies are discrete objects, so I write (Lacasa et al. 2013) :

$$n_{\text{gal}}(\mathbf{x}|i) = \sum_{j=1}^{N_{\text{gal}}(i)} \theta(\mathbf{x} - \mathbf{x}_j) \quad (3.43)$$

where j is the index of the random galaxies, $N_{\text{gal}}(i)$ is the –random– number of galaxies in the halo i , x_j is the –random– position of the j th galaxy, and $\theta(x)$ is the profile of a galaxy.

Accounting for this discreteness is necessary, as it allows to predict the shot-noise terms, and to do it self-consistently with the clustering terms. Eq.3.41 can then be rewritten as :

$$n_{\text{gal}}(\mathbf{x}) = \int dM d^3\mathbf{x}_h \sum_i \delta(M - M_i) \delta^{(3)}(\mathbf{x}_h - \mathbf{x}_i) \int d^3\mathbf{x}_g \sum_{j=1}^{N_{\text{gal}}(i)} \delta^{(3)}(\mathbf{x}_g - \mathbf{x}_j) \theta(\mathbf{x} - \mathbf{x}_j) \quad (3.44)$$

with M_i and \mathbf{x}_i the mass and position of the halo i .

This equation serves as the basis for the computation of the galaxy clustering throughout this thesis.

Within this formalism, we have :

$$\left\langle \sum_i \delta(M - M_i) \delta^{(3)}(\mathbf{x}_h - \mathbf{x}_i) \right\rangle = \frac{dn_h}{dM}(M) \quad (3.45)$$

which expresses the fact that the halo abundance is isotropic and given by the mass function.

Furthermore :

$$\langle \delta^{(3)}(\mathbf{x}_g - \mathbf{x}_j) \rangle_{\mathbf{x}_j} = u(\mathbf{x}_g - \mathbf{x}_i | M_i) \quad (3.46)$$

which expresses the fact that galaxy positions follow, on average, the halo profile.

I will further assume that galaxies are drawn independently in the halo.

Then galaxy correlation functions may be computed from Eq.3.44 ; for example the 2-point c.f. is :

$$\zeta_{\text{gal}}(|\mathbf{x} - \mathbf{x}'|) = \frac{\langle (n_{\text{gal}}(\mathbf{x}) - \bar{n}_{\text{gal}})(n_{\text{gal}}(\mathbf{x}') - \bar{n}_{\text{gal}}) \rangle}{\bar{n}_{\text{gal}}^2} \quad (3.47)$$

These correlation functions may be Fourier transformed to yield the polyspectra.

I have proposed in Lacasa et al. (2013) that these computations may be done through a diagrammatic method. It allows us to have a clear representation and an understanding of the different terms involved. It further allows us to avoid cumbersome calculations at high order, by replacing them with diagram drawings.

Note that, since Feynman diagrams (see e.g. Penco & Mauro 2006), diagrammatic method can be found in other contexts in physics and cosmology when perturbation theory is involved. However the diagrams proposed here are of a different nature as they do not involve perturbation theory. For example there are a finite number of them, for a given correlation function, and they yield an theoretically exact result (compared to a perturbative expansion where the result is known to be approximate).

For the polyspectrum of order n , the first step is to draw in diagrams all the possibilities of putting n galaxies in halo(s). Potentially, two or more galaxies can lie at the same point (“contracted”) for the shot-noise terms. Then for each diagram, the galaxies should be labeled, e.g., from 1 to n , as well as the haloes e.g. with α_1 to α_p .

Each diagram produces a polyspectrum term. This term contains a prefactor $1/\bar{n}_{\text{gal}}^n$ multiplied by an integral over the halo masses $\int dM_{\alpha_1 \dots \alpha_p}$ of several factors. The following “Feynman”-like rules prescribe these different factors :

- for each halo α_j there is a corresponding :
 - halo mass function $\frac{dn_h}{dM}|_{M_{\alpha_j}}$
 - average of the number of galaxy uplets in that halo.
e.g. $\langle N_{\text{gal}} \rangle$ for a single galaxy in that halo, $\langle N(N-1) \rangle$ for a pair etc.
 - as many halo profile $u(k|M_{\alpha_j})$ as *different* points, where $k = |\sum_{i \in \text{point}} \mathbf{k}_i|$.⁷
- the final factor is the halo polyspectrum of order p , conditioned to the masses of the corresponding haloes :

$$\mathcal{P}_{\text{halo}}^{(p)} \left(\sum_{i \in \alpha_1} \mathbf{k}_i, \dots, \sum_{i \in \alpha_p} \mathbf{k}_i \mid M_{\alpha_1}, \dots, M_{\alpha_p} \right)$$

where the sum $\sum_{i \in \alpha_j} \mathbf{k}_i$ runs over the indexes i of the galaxies inside the halo α_j .

Finally the possible permutations of the galaxy labels 1 to n in the diagram should be taken into account : the contribution is the sum over permutations which produce different diagrams. For example we will see in Sect.4.3.3 that some contributions to the bispectrum (namely 1-halo, 3-halo and shot1g) have a single term while others (namely 2-halo and shot2g) have 3 terms.

For illustration, these rules are applied in the case of the diagram represented in Fig.3.7.

This is a diagram involved in the computation of the trispectrum, it is rather complicated which allows to illustrate all the rules. The diagram involves two haloes labeled **a** and **b** and four galaxies labeled 1,2,3,4 with galaxies 3 and 4 being contracted. The latter means that the 4-point correlation function hits twice the same point with its last arguments ($\mathbf{x}_3 = \mathbf{x}_4$), so this is a shot-noise diagram. This diagram yields the following term for the trispectrum :

$$T(\vec{k}_1, \dots, \vec{k}_4, z) = \frac{1}{\bar{n}_{\text{gal}}(z)^4} \int dM_{\text{ab}} \frac{dn_h}{dM} \Big|_{M_{\text{a}}} \frac{dn_h}{dM} \Big|_{M_{\text{b}}} \langle N_{\text{gal}}(M_{\text{a}}) \rangle \langle N_{\text{gal}}(M_{\text{b}})(N_{\text{gal}}(M_{\text{b}}) - 1) \rangle \\ \times u(k_1|M_{\text{a}})u(k_2|M_{\text{b}})u(|\vec{k}_3 + \vec{k}_4||M_{\text{b}}) P_{\text{halo}}(k_1|M_{\text{a}}, M_{\text{b}}) \quad (3.48)$$

The **prefactor** is dictated by the first “Feynman”-like rule, then we have an integral over the mass of both haloes **a** and **b**, their **mass functions**, their **galaxy occupation** (with an average number of duplets $N(N-1)$ in halo **b** as there are two different galaxies), their **profile**, and finally the polyspectrum linking the haloes, here the **halo power spectrum** conditioned to the respective halo masses.

⁷For example $k = k_i$ for a non-contracted galaxy i , while $k = |\mathbf{k}_{i_1} + \dots + \mathbf{k}_{i_q}|$ for a galaxy contracted q times with labels $i_1 \dots i_q$.

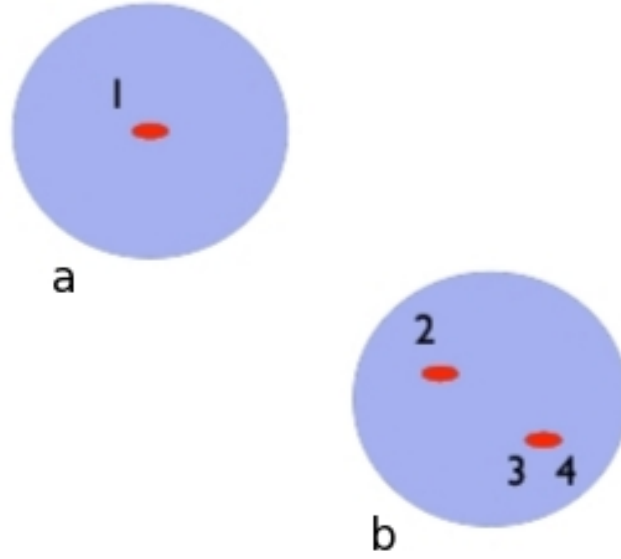


Figure 3.7: A diagram involved in the computation of the galaxy trispectrum : 2-halo 3-galaxies with contraction in the second halo.

I have described in the previous sections the first elements (mass function, HOD, halo profile), the last elements needed are the halo polyspectra. They are biased versions of the dark matter polyspectra through deterministic biasing discussed in Sect.3.4.2⁸. At high order they have several possible origins : the first order biasing of the corresponding dark matter polyspectrum, either primordial (primordial non-Gaussianity) or from perturbation theory, or the higher order biasing of lower order dark matter polyspectra. This will be discussed in Sect.4.3.3 when this formalism will be applied to the CIB (see also Lacasa et al. 2013; Pénin et al. 2013).

A concluding remark for this section, is that we can resum shot-noise diagrams. For example, it is easily seen that a diagram for the n -th order polyspectrum with a contraction of order 2 corresponds to a diagram for the $(n-1)$ th order polyspectrum without contraction. And the polyspectrum terms that these diagrams yield differ only by a factor $\frac{1}{\bar{n}_{\text{gal}}(z)}$ and the different arguments \mathbf{k}_i . Hence, when we sum all these diagrams with a contraction of order 2 we get the lower order polyspectrum :

$$\mathcal{P}_{\text{gal,contrac2}}^{(n)}(\mathbf{k}_{1\dots n}, z) = \frac{1}{\bar{n}_{\text{gal}}(z)} \mathcal{P}_{\text{gal,non-shot}}^{(n-1)}(\mathbf{k}_1 + \mathbf{k}_2, \mathbf{k}_{3\dots n}, z) + \text{perm.} \quad (3.49)$$

where there are $\binom{2}{n}$ permutation terms, and the $(n-1)$ th order polyspectrum only contains non shot-noise terms, as the n -th order diagrams considered contain strictly a contraction of order 2 and no other contractions.

⁸As already pointed out in Sect.3.4.2, this description of the halo polyspectra breaks down on small (non-linear) scale. However on these scales, the galaxy polyspectrum will be dominated by terms involving halo profiles and not halo correlations. For example at the power spectrum level, the description of the 2-halo term breaks down at high k , but this is, as it is meant to be, when the 1-halo and shot-noise terms become dominant.

One can further develop resummation, which is of interest at high orders, but for the illustrations I will present in Sect.4.3.3, Eq.3.49 is sufficient.

Hence I have presented a diagrammatic method to compute galaxy polyspectra which brings us the power and simplicity of drawings to compute otherwise cumbersome equations at high order. Furthermore, this diagrammatic method can be easily adapted to model complexifications (different galaxy populations, additional galaxy and/or halo properties etc), as will be discussed in the perspectives.

In this chapter, I have described the very basis of the current cosmological description of the universe. I have described the inflation paradigm and how it provides a model producing the primeval cosmological perturbations. I have shown how these perturbations then evolve to produce the two main observables of modern cosmology : the Cosmic Microwave Background and the Large Scale Structure. Both observables can be described as random fields and are hence naturally fit for statistical characterisation with the tools developed in the previous chapters. In particular, I have shown how non-linearity of the physics generates non-Gaussian signatures in both these observables. Finally, I have presented a diagrammatic formalism allowing the prediction of high order polyspectra for a standard tracer of the Large Scale structure, the galaxy density field, and these results will be further used in the following chapter to describe the Cosmic Infrared Background.

Chapter 4

Extragalactic foregrounds : tracers of the Large Scale Structure

Contents

4.1	Foregrounds : introduction	67
4.2	Sunyaev-Zel'dovich effect	71
4.2.1	Introduction : ionisation and photon scattering	71
4.2.2	Description with the halo model	72
4.3	Extragalactic point-sources	74
4.3.1	Introduction	74
4.3.2	Prescription for the non-Gaussianity of clustered and unclustered sources	78
4.3.3	CIB description with the halo model	81

This chapter introduces the different signals that the Large Scale Structure produces at CMB frequencies. It focuses particularly on two of them, tracing respectively the ionised gas and the dust in the universe. I show how they can be described using tools and formalism introduced in the previous chapters, and particularly how their non-Gaussianity can be modeled in order to constrain their respective models.

4.1 Foregrounds : introduction

The CMB follows a blackbody spectrum which peaks around $1 \text{ mm} \approx 300 \text{ GHz}$. However, the CMB is not the only signal in the sky. As shown in Fig.4.1, the CMB clearly dominates the total power at millimeter wavelengths, while the radio/infrared background becomes dominant respectively at smaller/larger frequencies.

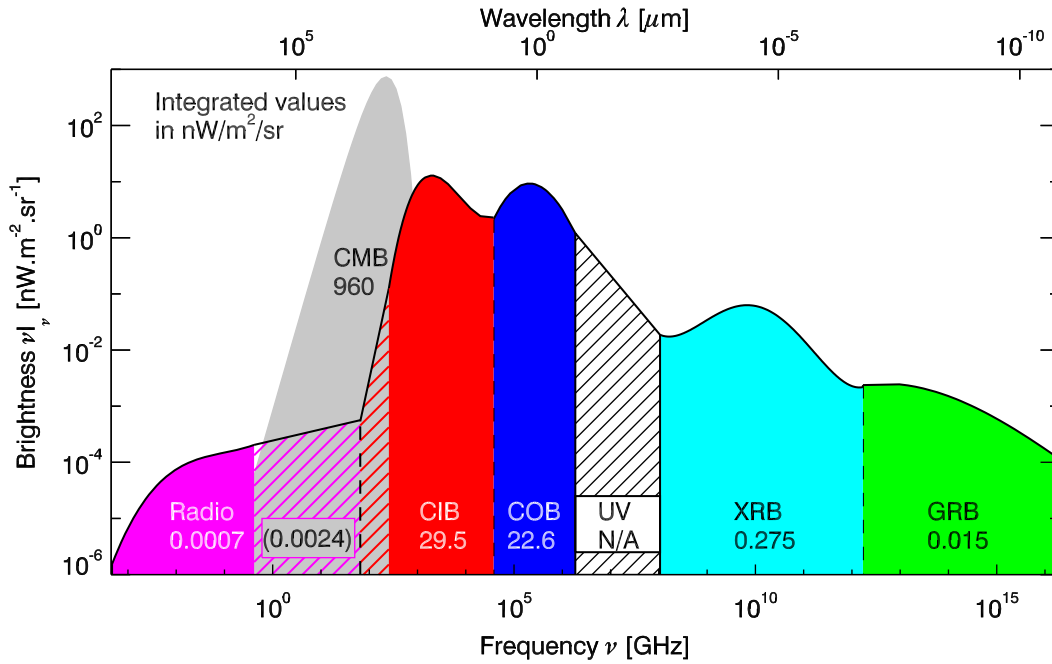


Figure 4.1: Extragalactic background light energy density as a function of frequency. Courtesy of Hervé Dole. From left to right : radio background, Cosmic Microwave Background, Cosmic Infrared Background, Cosmic Optical Background, UltraViolet background, X-Ray Background and Gamma Ray Background.

Although the CMB dominates in terms of total power, it is remarkably uniform over the sky. As stated before, its primary anisotropies only appear at the 10^{-5} level. Observationally the first detectable anisotropy, at the 10^{-3} level, is the so-called kinetic dipole due to the motion of the observer with respect to the surface where CMB photons were emitted (last scattering surface hereafter)¹. When this dipole is removed, the main feature visible on the sky is the Milky way, as can be seen for example on the first *Planck* released sky image Fig.4.2.

¹At first order in $\beta = \frac{v_{\text{obs}}}{c}$ this effect creates a dipolar anisotropy, but also a whole aberration of the sky ; it furthermore yields a kinetic quadrupole and monopole at second order in β

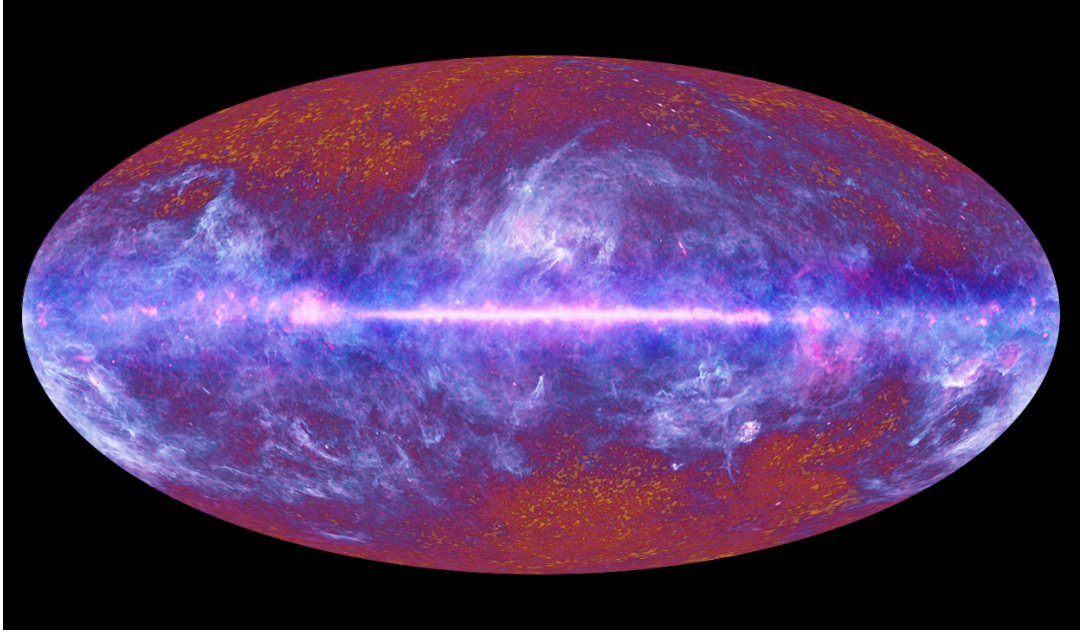


Figure 4.2: Multifrequency image of the microwave sky by *Planck* (ESA/ LFI & HFI Consortia, July 2010 press release). The galactic emission, mostly in blue, is clearly visible at the equator in this projection. It obscurs the CMB, which remains visible at high galactic latitudes.

There are four main processes of galactic emission :

- **Synchrotron** : the Bremsstrahlung emission of relativistic electrons spiraling in the galactic magnetic field. This emission peaks in the radio domain and decreases with frequency with a steep slope (see Davies & Wilkinson 1998, for a review).
- **Free-free** : emission generated when free electrons are scattered by ions, e.g. in an ionised hydrogen cloud. This emission peaks in the radio domain and decreases with frequency with a flatter slope (see Bartlett & Amram 1998, for a review).
- **Thermal dust** : thermal emission from cool dust, typically at a temperature of 20 K. This emission follows a greybody law and peaks in the infrared domain (see Draine 2004, for a review).
- **Anomalous microwave emission (AME)** : only detected in the last decade, and thought to be produced by spinning dust grains. It emits in the radio domain with a peak typically between 20 GHz and 40 GHz (see Kogut 1999, for a review).

Other galaxies in the universe also emit through these processes, and may be visible outside of the galactic plane. Their emission, redshifted by the expansion, produces an extragalactic background light which traces the galaxy distribution. Galaxies are most often dominated by one emission type and are thus broadly classified into two populations :

- **radio sources** are dominated by synchrotron and/or free-free emission. They correspond to galaxies with a central Active Galactic Nucleus (AGN) (see Toffolatti et al. 1999, for a review).

- **infrared sources** are dominated by thermal dust emission. They correspond to dusty star-forming galaxies and form the Cosmic Infrared Background (CIB) (see Gispert et al. 2000; Lagache et al. 2005, for reviews).

Both populations will be described in more details in Sect.4.3.1.

The brightest of these sources are detected and characterised, the latest study being by *Planck* in Planck Collaboration XXVIII (2013). The study of the unresolved sources and their statistics has been an important part of my work.

Further anisotropies, called secondary, are present at CMB frequencies. They do not come from an additional emission process. Instead, they are distortions of the CMB emission, along its path from the last scattering surface to the observer. Three main distortions produce the most observable signals :

- **Lensing** : the trajectory of photons may deviate from a straight line, when they encounter gravitational potentials which produce a local spatial curvature. The angular scales of interest for the CMB do not probe magnification or multiple-imaging induced by strong lensing. Instead there is a distortion by the potentials on large scale, in the regime of weak lensing. The net effect is to remap the primordial anisotropies, i.e. $T_{\text{obs}}[\hat{n}] = T_{\text{prim}}[\hat{n}'(\hat{n})]$.

A review of the theory and observation of the CMB lensing is given by Hanson et al. (2010), and the latest observations by Planck Collaboration XVII (2013) have produced the first estimated lensing map.

- **integrated Sachs-Wolfe effect (iSW)** : if the photons cross gravitational potentials that are evolving, the photons experience an energy shift. For example if the potential inhomogeneities flatten, a photon gains more energy falling down a well than it loses energy climbing up the exit. In a matter-dominated euclidean universe, gravitational potentials are however constant in the linear regime. Hence the iSW effect happens either when the dynamic differs from matter-dominated (e.g., when dark energy becomes important) or through non-linear gravitational infall (sometimes called Rees-Sciama effect (Rees & Sciama 1968) or non-linear iSW). The iSW effect due to dark energy (sometimes called late iSW) peaks on large angular scales, and much more information about it can be found in the PhD thesis by Ilic (2013).

A review of the theory and observation of the iSW effect is given by Giannantonio (2010), and it has been detected and characterised with the latest observations by Planck Collaboration XIX (2013).

- **Sunyaev-Zel'dovich effect (SZ)** : CMB photons may also be scattered by free electrons through Compton scattering. The thermal energy of these electrons creates y -type spectral distortions of the CMB blackbody named the thermal Sunyaev-Zel'dovich (tSZ) effect. Furthermore, the bulk velocity of these electrons creates a temperature anisotropy, named the kinetic Sunyaev-Zel'dovich (kSZ) effect, which depends on the orientation of the velocity with respect to the line of sight.

The tSZ effect will be described in more details in Sect.4.2. Theory and observations reviews of the SZ effect are given by Cooray et al. (2005) and Bartlett (2004), and the latest observations *Planck* have produced the largest SZ-based cluster catalogue (Planck Collaboration XXIX 2013) and the first estimated full-sky SZ map (Planck Collaboration XXI 2013).

- Note that there are other processes leaving secondary imprint on the CMB, such as the Ostriker-Vishniac effect (Ostriker & Vishniac 1986) or inhomogeneous reionisation (Aghanim et al. 1996). They however have a much lower amplitude and peak at small angular scales. See Aghanim et al. (2008) for a thorough and complete review of the sources of secondary anisotropies.

Note that, as they trace the Large Scale Structure at low redshift ($z = 0 - 1$), the secondary anisotropies are significantly correlated together and have a non-Gaussian distribution. As already mentioned, the iSW-lensing bispectrum is of particular importance for CMB NG studies. Indeed the large scale iSW signal modulates the small scale lensing power, and hence the bispectrum peaks in the squeezed limit and is a contaminant for local type primordial NG estimation (at the level $\Delta f_{\text{NL}} \sim 10$, see Serra & Cooray (2008); Kim et al. (2013)).

Note also that the first three listed phenomena can potentially also affect photons that are produced by sources other than the CMB (as described previously : radio sources, CIB etc). However these latter signals were emitted much later on, so that along their way they did not cross appreciable gravitational potentials or free electron densities. Thus the distortion of foreground signals is mostly negligible, except for the occasional strong lensing of an extragalactic point-source, which magnifies its flux.

Simulations of these foregrounds have been produced, e.g. by CMB experiment teams to assess the foreground contamination to CMB analysis. In particular, I have been using two sets of simulations respectively aimed for the Atacama Cosmology Telescope (ACT, Kosowsky 2003) and *Planck* experiments (Planck Collaboration et al. 2011d).

The first set is composed of all-sky simulated maps by Sehgal et al. (2010) at 30, 90, 148, 219 and 350 GHz, and publicly available². I have been using the simulated maps of the IR and radio point-sources and of the thermal Sunyaev-Zel'dovich effect. The maps are based on N-body simulations of the large scale structure, with a volume $1000 h^{-1}\text{Mpc}$ on a side, produced using a tree-particle mesh code. The simulated octant of the sky was replicated, which produces spurious effects on large scales, particularly up to the octopole $\ell = 3$; however I have avoided these scales for my studies. Furthermore, the simulations reproduce the observables available at the time of their productions. Indeed for the CIB, they match the upper limit on the total intensity by Fixsen et al. (1998), the dust parameters by Knox et al. (2004); Dunne et al. (2000), the source counts by Coppin et al. (2006); Austermann et al. (2010); Patanchon et al. (2009) and the constraint on the power spectrum by Lagache et al. (2007); Viero et al. (2009); Reichardt et al. (2009). However they do not reproduce the scale dependence of the later measurements of the power spectrum (Planck-Collaboration 2011). For radio sources, the simulations are based on a radio model consistent with observations (Lin et al. 2009, 2010) and reproduce the radio luminosity function (Laing et al. 1983), they however do not reproduce exactly the number counts found in the later results by Planck Collaboration et al. (2011b); Planck Collaboration XXVIII (2013).

The second set of simulations is the *Planck* sky model (PSM hereafter, Delabrouille et al. 2013) and consists of simulations of all the Galactic and extragalactic foregrounds to the CMB at all the *Planck* frequencies. That is, at 30-44-70 GHz for the Low Frequency Instrument (LFI) and at 100-143-217-353-545-857 GHz for the High Frequency

²http://lambda.gsfc.nasa.gov/toolbox/tb_cmbsim_ov.cfm

Instrument (HFI). Note that the IR and radio point-sources simulations are not based on a large scale structure simulation but partly on observations and partly on random realisations. The PSM software is continuously evolving to reproduce observations, in particular *Planck* results. I have used the PSM simulations in particular for the *Planck* tSZ analysis, see Sect.5.3.

4.2 Sunyaev-Zel'dovich effect

4.2.1 Introduction : ionisation and photon scattering

Energetic free electrons are found in galaxy clusters, and as they are not distributed uniformly they leave secondary anisotropies in the CMB. Indeed Sunyaev & Zeldovich (1970, 1972) have shown that thermal electrons inject energy in the CMB photons through inverse Compton scattering. Through this scattering, the number of photons is conserved but they are transferred to higher frequencies. The result is a distortion of the spectrum, with a depletion at low frequencies and an excess at high frequencies, compared to the CMB blackbody spectrum. In the case of optically thin clusters (on average photons experience a single scattering) and non-relativistic electrons, the spectral distortion has a universal form and is characterised by a single amplitude parameter, the Comptonisation parameter y :

$$y = \int n_e \sigma_T \frac{k_B T_e}{m_e c^2} dl \quad (4.1)$$

where n_e and T_e are respectively the electron density and temperature in the cluster, σ_T is the Thomson cross-section, m_e is the electron rest-frame mass, and the integral runs over the line-of-sight.

A y -type distortion of the CMB spectrum is shown in Fig.4.3 for $y = 0.1$, although the effect is much more subtle for real clusters ($y_{\text{rms}} \sim 10^{-6}$).

The bulk velocity of the electrons in the clusters also leaves an imprint on the CMB photons. Indeed, the electrons tend to define a frame that the photons should be at rest with respect to. When this frame is moving with respect to the observer, the photons experience a Doppler redshift (outward velocity) or blueshift (inward velocity). This is the same process that leads to Doppler primary anisotropies of the CMB, due to the velocity of the plasma at the last scattering surface. The effect is achromatic and hence photons keep a blackbody spectrum albeit with a relative temperature change :

$$\frac{\Delta T}{T_{\text{CMB}}} = \tau \frac{\vec{v}_e \cdot \hat{n}}{c} \quad (4.2)$$

where τ is the optical depth of the cluster. Fig.4.3 shows this kinetic SZ effect in the case of a receding cluster.

Observationally, the tSZ effect has been detected for a long time (e.g., Wilbanks et al. 1994; Lamarre et al. 1998), and recently *Planck* has produced the first estimated map of the tSZ effect over the sky (Planck Collaboration XXI 2013). But the detection of the kSZ effect is more challenging, due to its low amplitude and to its spectrum being indistinguishable from a legitimate CMB anisotropy. Its statistical detection over thousands of clusters has nonetheless been recently claimed by Hand et al. (2012).

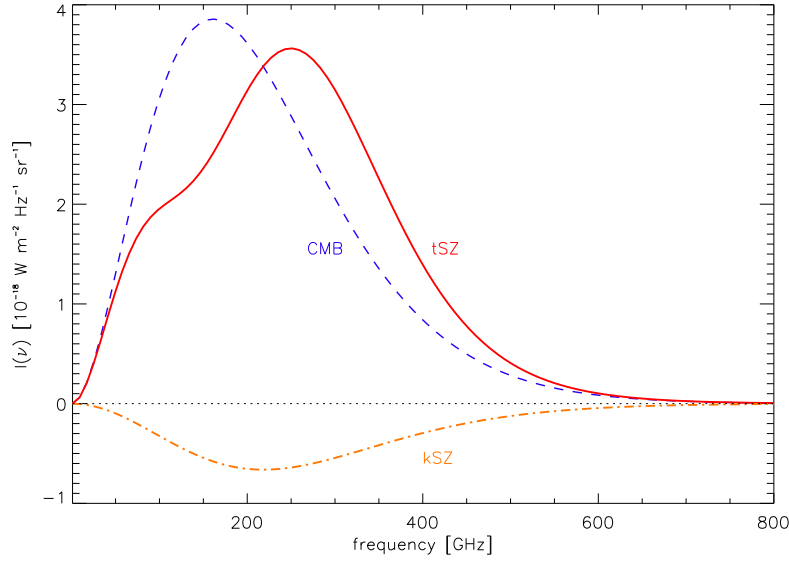


Figure 4.3: Emission law for the primordial CMB (blue), for CMB photons distorted by a thermal SZ effect (red), and by the kinetic SZ effect (orange, $\Delta I(\nu)$) in the case of a recessing cluster.

4.2.2 Description with the halo model

If we are interested in the spatial variations of the thermal SZ signal, they can be described with the halo model previously introduced. Indeed the 3D Comptonisation parameter reads :

$$y(\mathbf{x}, z) = \int dM d^3\mathbf{x}_h y(M, z) u(\mathbf{x} - \mathbf{x}_h | M, z) \sum_i \delta(M - M_i) \delta^{(3)}(\mathbf{x}_h - \mathbf{x}_i) \quad (4.3)$$

where I splitted the y-profile into the total flux of the halo $y(M, z)$ and its normalised profile $u(\mathbf{x} | M, z)$ in analogy with the CIB case (see Sect.4.3.3).

Thus the Fourier transform of $y(\mathbf{x}, z)$ is :

$$y(\mathbf{k}, z) = \int dM d^3\mathbf{x}_h y(M, z) u(k | M, z) e^{-i\mathbf{k} \cdot \mathbf{x}_h} \sum_i \delta(M - M_i) \delta^{(3)}(\mathbf{x}_h - \mathbf{x}_i) \quad (4.4)$$

where we assumed sphericity of the halos³ so that $u(\mathbf{k} | M, z) = u(k | M, z)$.

The observed Comptonisation parameter \tilde{y} is the line-of-sight integral of $y(\mathbf{x}, z)$:

$$\tilde{y}(\hat{n}) = \int dz \frac{dV_c}{dz d\Omega} y(r\hat{n}, z)$$

Hence, through Sect.2.4.4, if the 3D polyspectrum is diagonal-independent and with Limber's approximation, the angular polyspectrum reads :

$$\mathcal{P}_{\text{tSZ}}^{(n)}(\ell_{1\dots n}) = \int \frac{dz}{r^{2n-2}} \frac{dr}{dz} a^n(z) \mathcal{P}_{\text{tSZ,3D}}^{(n)}(\mathbf{k}_{1\dots n}^*, z) \quad (4.5)$$

³Indeed the halo shape has been shown to have a small ($\leq 10\%$) effect on the dark matter bispectrum (Smith et al. 2006).

with as usual $k_i^* = \ell_i/r(z)$.

The 3D polyspectra can be computed through the connected correlation functions of Eq.4.4. For illustration, the computation of the power spectrum is :

$$\begin{aligned}
(2\pi)^3 \delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2) P_{\text{tSZ}}(k_1, z) &= \langle y(\mathbf{k}_1, z) y(\mathbf{k}_2, z) \rangle \\
&= \int \left[dM_i d^3 \mathbf{x}_i^h y(M_i, z) u(k_i|M_i, z) e^{-i\mathbf{k}_i \cdot \mathbf{x}_i^h} \right]_{i=12} \\
&\times \left\langle \sum_{i_1 i_2} \delta(M_1 - M_{i_1}) \delta^{(3)}(\mathbf{x}_1^h - \mathbf{x}_{i_1}) \delta(M_2 - M_{i_2}) \delta^{(3)}(\mathbf{x}_2^h - \mathbf{x}_{i_2}) \right\rangle
\end{aligned} \tag{4.6}$$

It splits into two terms : $P(k) = P^{1h}(k) + P^{2h}(k)$, as the sum $\sum_{i_1 i_2}$ is splitted. The first term is the 1-halo term, it corresponds to $\sum_{i_1=i_2}$, while the second term is the 2-halo term and corresponds to $\sum_{i_1 \neq i_2}$ (Cooray & Sheth 2002; Ma & Fry 2002). We find :

$$P_{\text{tSZ}}^{1h}(k, z) = \int dM \frac{dn_h}{dM} [y(M, z) u(k|M, z)]^2 \tag{4.7}$$

and :

$$P_{\text{tSZ}}^{2h}(k, z) = \int dM_{12} \left[y(M_i, z) u(k|M_i, z) \frac{dn_h}{dM} \Big|_{M_i} \right]_{i=12} P_{\text{halo}}(k|M_1, M_2, z) \tag{4.8}$$

Correspondingly the angular power spectrum splits into 1-halo and 2-halo terms :

$$C_\ell^{\text{tSZ}} = C_\ell^{1h} + C_\ell^{2h} \tag{4.9}$$

Fig.4.4 shows the first measurement of the tSZ angular power spectrum by *Planck*, and the best-fit halo model from Planck Collaboration XXI (2013), fitting for the cosmological parameters σ_8 and Ω_m which have the strongest influence. Because the tSZ signal is dominated by massive low-redshift clusters, the 1-halo term completely dominates the power spectrum, except marginally on the very first multipoles. (Note that this may change if applying a mass cut, i.e. masking detected clusters above a certain mass)

I also derive here the equations for the tSZ bispectrum. The 3-point correlation function of Eq.4.4 splits into three terms, depending on the splitting of the sum over halo indices $\sum_{i_1 i_2 i_3}$. First, the 1-halo term corresponds to $\sum_{i_1=i_2=i_3}$, and after some computations it writes :

$$B_{\text{tSZ}}^{1h}(k_{123}, z) = \int dM \frac{dn_h}{dM} y(M, z)^3 u(k_1|M, z) u(k_2|M, z) u(k_3|M, z) \tag{4.10}$$

The 2-halo term corresponds to $\sum_{i_1=i_2 \neq i_3} + 2$ permutations (perm. hereafter) and reads :

$$\begin{aligned}
B_{\text{tSZ}}^{2h}(k_{123}, z) &= \int dM_{\text{ab}} \frac{dn_h}{dM} \Big|_{M_{\text{a}}} \frac{dn_h}{dM} \Big|_{M_{\text{b}}} y(M_{\text{a}}, z)^2 y(M_{\text{b}}, z) u(k_1|M_{\text{a}}, z) \\
&\times u(k_2|M_{\text{a}}, z) u(k_3|M_{\text{b}}, z) P_{\text{halo}}(k_3|M_{\text{a}}, M_{\text{b}}, z) + 2 \text{ perm.}
\end{aligned} \tag{4.11}$$

Finally, the 3-halo term corresponds to $\sum_{i_1 \neq i_2 \neq i_3}$, and writes :

$$B_{\text{tSZ}}^{3h}(k_{123}, z) = \int dM_{123} \left[y(M_i, z) u(k_i|M_i, z) \frac{dn_h}{dM} \Big|_{M_i} \right]_{i=123} B_{\text{halo}}(k_{123}|M_{123}, z) \tag{4.12}$$

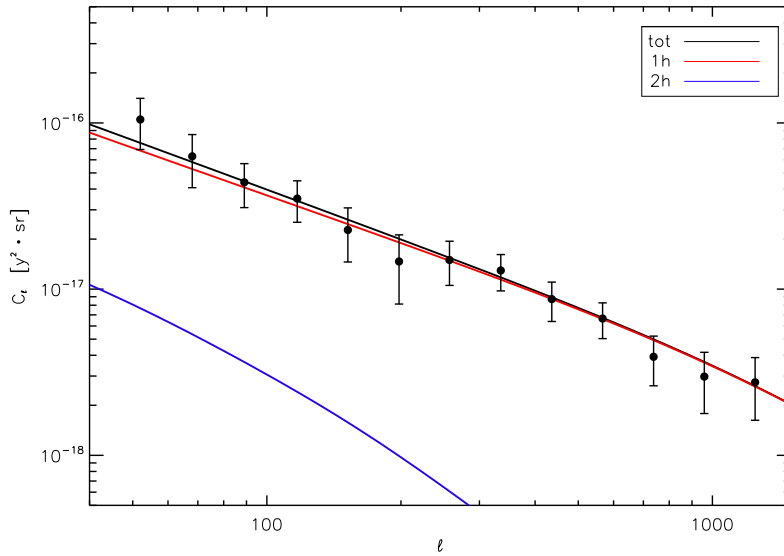


Figure 4.4: tSZ angular power spectrum measured with *Planck* (Planck Collaboration XXI 2013, black points with error bars) and the best-fit power spectrum with the halo model (black curve). The red and blue curve show respectively the 1-halo and 2-halo term of the best-fit model.

I do not enter into further details in this section –in particular on the halo power spectrum and bispectrum–, as most of the issues will be treated in Sect.4.3.3 for the CIB with similar equations.

Because the tSZ signal is dominated by massive low-redshift clusters, I expect high order polyspectra to be dominated by the 1-halo term as for the power spectrum, except possibly when low multipoles are involved (e.g. in the squeezed configurations for the bispectrum). Measurement performed in Sect.5.3 will be shown to support this view. Moreover, because it is dominated by massive low-redshift clusters, the tSZ signal is highly non-Gaussian (Komatsu & Seljak 2002). This will also be shown observationally in Sect.5.3, and motivates the use of non-Gaussianity studies of the tSZ observations to extract more statistical information, and hence more powerful model constraints.

4.3 Extragalactic point-sources

4.3.1 Introduction

As already discussed in Sect.4.1, the galaxies' emission at microwave frequencies yields a foreground to the CMB. They are classified as extragalactic point-sources as, with the resolution of microwave experiments, their angular size is much smaller than the beam (except for the nearest galaxies, such as Andromeda).

Radio sources include galaxies with a strong radio emission due to synchrotron and/or free-free processes. In these galaxies an Active Galactic Nucleus, most probably resulting from a super-massive black hole, produces a magnetised jet which ejects ionised matter out of the galactic core. The strong magnetic field in the core and the jet produces

synchrotron emission, while the lobes of the jet (shock front) emit through free-free process. Depending on the orientation of the jet with respect to the observer, on the temperature of the lobes, on the velocity of electrons in the jet, on the absorption by the galactic disk etc, the observed flux may be dominated either by synchrotron or free-free, or be a mix, and be more or less powerful. The typical Spectral Energy Distribution (SED) of a radio source is steep ($I(\nu) \propto \nu^{-\alpha}$ with $\alpha \sim 0.7$, see e.g. Marscher 1996; Tucci et al. 2011), hence few of them are visible at high frequencies.

At microwave frequencies the detectable radio sources are the brightest ones, with a few hundreds being detected by *WMAP* (Wright et al. 2009) and about a thousand detected by *Planck* (Planck Collaboration XXVIII 2013), and they dominate the extragalactic radio emission (Toffolatti et al. 1998). They reside at somewhat low redshifts (peak at $z \sim 1$, see Tucci et al. 2011), and can be considered as randomly distributed over the sky (Toffolatti et al. 1998; González-Nuevo et al. 2005). Hence they form a white noise entirely characterised by its 1-point p.d.f., i.e., the number counts $\frac{dn}{dS}$. Thus, their polyspectra are constant and given by :

$$\mathcal{P}_{\text{RAD}}^{(n)} = \int S^n \frac{dn}{dS} dS \quad (4.13)$$

which stems from the properties of cumulants and the fact that the sources are independent. Let us note that the integral runs from $S = 0$ to S_{cut} . The flux cut, S_{cut} , is the flux of the faintest detected source, i.e. all sources below that flux are present in the map and all brighter sources are assumed to be masked⁴. Indeed, for the study of signal other than that of radio sources, one wants to minimise the radio contamination.

In particular, the bispectrum of unresolved radio-sources is traditionally noted b_{PS} , with :

$$b_{\text{PS}} = \int_0^{S_{\text{cut}}} S^3 \frac{dn}{dS} dS \quad (4.14)$$

This is a convention dating back to *WMAP* (Komatsu et al. 2005), as radio sources are the only contributing sources at *WMAP* frequencies.

During my work, I have been using number counts from the radio sources model by Tucci et al. (2011), which is the most recent one and reproduces the observed number counts. Fig.4.5 shows these number counts at 30 GHz, the lowest *Planck* frequency.

In the figure I plot $S^{5/2} \frac{dn}{dS}$, this is the so-called Euclidean normalised convention. The reason for this convention is that in a flat static universe (Euclidean) the number counts go as $\frac{dn}{dS} \propto S^{-5/2}$. As a matter of fact, if we consider galaxies with a given luminosity L uniformly distributed with density ρ_L , an observer sees $n(> S) = \rho_L \times \frac{4}{3}\pi R^3$ sources above flux S , with $S = \frac{L}{4\pi R^2}$. Hence $n(> S) \propto S^{-3/2}$ and derivation yields $\frac{dn}{dS} \propto S^{-5/2}$. This stands for all L , so we can integrate over L which does not change the functional form but simply the proportionality constant.

Thus any departure from this behaviour comes from expansion and/or redshift evolution of sources. In particular the decrease in Fig.4.5 when $S \rightarrow 0$ comes from the fact that

⁴The situation is often more complex than a flux cut consideration, and requires a selection function which expresses the estimated detection probability depending on the flux. This may happen for example when the noise and/or contaminants are anisotropic. I will not enter into these considerations here, let us just note that the flux cut case is the special case where the selection function is a Heaviside step function.

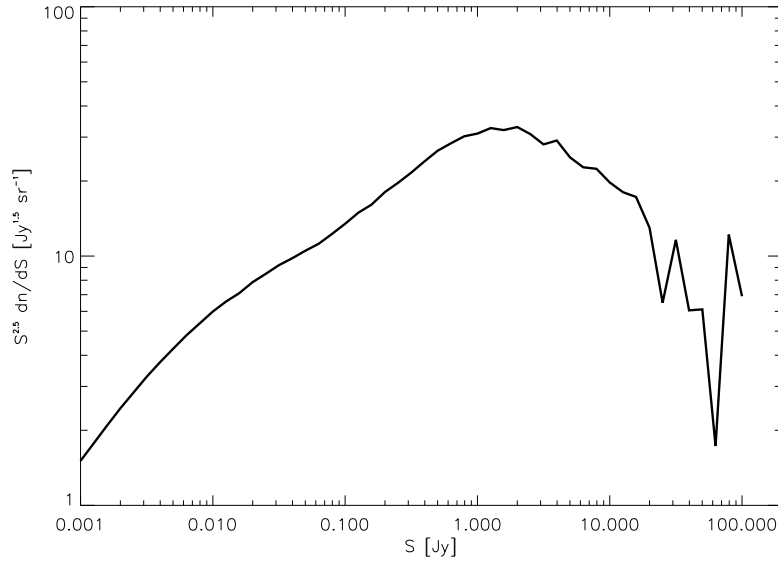


Figure 4.5: Number counts for radio sources at 30 GHz with the model by Tucci et al. (2011), in the Euclidean normalisation (see text for details).

sources started emitting a finite time ago, so that there are no extremely distant and faint sources.

Infrared (IR) sources are galaxies with a strong emission from thermal dust. This dust emits with a greybody law at a temperature between 10 and 30 K, heated by the UV emission from young stars (Gispert et al. 2000). These sources are hence called dusty star-forming galaxies (DSFG), and, compared to radio sources, they reside at higher redshifts : the star formation peaks towards $z \sim 2$ (Lagache et al. 2005).

IR sources have been characterised at the number counts level, and statistical methods have been proposed to try to constrain the number counts below the detection limit, e.g. with so-called $P(D)$ analysis (Patanchon et al. 2009) or with stacking (B  thermin & Dole 2011). A few thousands of them have been detected individually with *Planck* higher frequencies channels (Planck Collaboration et al. 2011b; Planck Collaboration XXVIII 2013), but there are in fact countless unresolved sources in a *Planck* beam⁵. IR sources are strongly clustered in dark matter halos, so that their spatial distribution yields a non-trivial power spectrum. In the latest years, the CIB power spectrum has been more and more precisely measured by different experiments : BLAST (Viero et al. 2009), Spitzer (Kashlinsky et al. 2012), Herschel (Amblard et al. 2011; Thacker et al. 2013) and *Planck* (Planck-Collaboration 2011) for the most recent. Fig.4.6 by Thacker et al. (2013) summarises the current state of the art. New measurements of this power spectrum by *Planck* will soon become public, I describe in Sect.5.2.3 my measurement of the CIB bispectrum.

The CIB power spectrum can be modeled with the halo model, and it splits into three terms : 2-halo, 1-halo and shot-noise. I will describe more precisely how these

⁵For example, using the number counts by B  thermin & Dole (2011), at 545 GHz there are a thousand sources between 100 m Jy and the flux cut of the Early Release Compact Source Catalog (Planck Collaboration et al. 2011b) in a *Planck* beam of 4.72 arcmin.

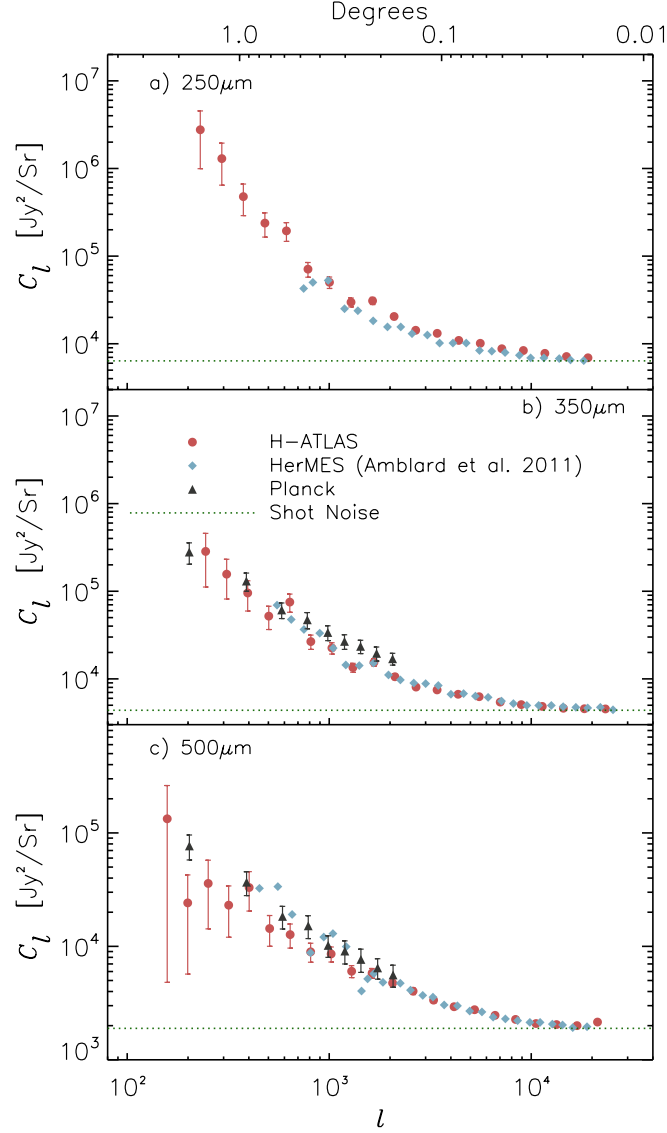


Figure 4.6: Different measurements of the CIB power spectrum at 1200, 857 and 550 GHz (from top to bottom). Figure by Thacker et al. (2013). The *Planck* data points are from the early results (Planck-Collaboration 2011).

terms arise in Sect.4.3.3. Let us note however that the 2-halo term dominates on large angular scales, then the 1-halo term dominates and finally the shot-noise dominates on the smallest angular scales. For the *Planck* resolution, the shot-noise is subdominant over the whole range of multipoles of interest, as can be seen in Fig.4.6. The shot-noise is a term coming from the discreteness of galaxies and is independent of their clustering ; it corresponds to Eq.4.13 for $n = 2$ and with IR number counts. In fact, in the case of radio sources, the shot-noise is the only contribution to the polyspectra as their clustering is negligible.

As the clustering of IR sources is dominated by faint unresolved sources, the flux cut has few effect on the CIB power spectrum. It mostly affects the shot-noise as the latter is sensitive to bright sources, but the effect on the clustering terms is small. More generally,

as the CIB is mostly unresolved at CMB frequencies, it is the stealthiest contaminant for most CMB studies ; it is taken special care of e.g. for primary CMB and cosmological parameter constraints (e.g., Planck Collaboration XVI 2013) but also for the tSZ (Planck Collaboration XXI 2013), lensing studies (Planck Collaboration XVIII 2013) etc.

4.3.2 Prescription for the non-Gaussianity of clustered and un-clustered sources

This section is based on results published in Lacasa et al. (2012), using the previously described simulations by Sehgal et al. (2010).

I have proposed a phenomenological prescription for the distribution of clustered sources, based on their number counts and power spectrum, following pioneering work by Argüeso et al. (2003). Namely, a map of a population of clustered sources is a white-noise drawn from the number counts, which has then been convolved by the power spectrum. By construction this prescription reproduces the 1- and 2- point correlation functions.

At the bispectrum level, the prescription yields :

$$b_{123}^{\text{clust}} = \alpha \sqrt{C_{\ell_1}^{\text{clust}} C_{\ell_2}^{\text{clust}} C_{\ell_3}^{\text{clust}}} \quad (4.15)$$

with :

$$\alpha = \frac{b_{123}^{\text{white}}}{\sqrt{C_{\ell_1}^{\text{white}} C_{\ell_2}^{\text{white}} C_{\ell_3}^{\text{white}}}} = \frac{\int S^3 \frac{dn}{dS} dS}{\left(\int S^2 \frac{dn}{dS} dS \right)^{3/2}} \quad (4.16)$$

Note that if the sources are not correlated, the prescription reduces to $b_{123} = \int S^3 \frac{dn}{dS} dS$, i.e. the shot-noise value.

If we have two populations of sources contributing to the same frequency, one clustered and one unclustered, as is the case with IR and radio sources, we assume that the populations are independent so that their bispectra add up :

$$b_{123} = b^{\text{RAD}} + \alpha \sqrt{C_{\ell_1}^{\text{IR}} C_{\ell_2}^{\text{IR}} C_{\ell_3}^{\text{IR}}} \quad (4.17)$$

I have developed a tool to compute the bispectrum of full-sky maps in the HEALPix format (Górski et al. 2005). I computed the bispectra of the simulated maps by Sehgal et al. (2010) at each frequency, for radio sources, infrared sources and their combination. We indeed see that the radio bispectrum is flat at all frequencies, whereas the IR bispectrum strongly depends on the multipoles, with orders of magnitude difference between large scales and small scales. For illustration I show in Fig.4.7 the bispectrum of the superposition of radio and IR maps in the equilateral configuration ($b_{\ell\ell\ell}$) depending on frequency.

At low frequencies, 30 & 90 GHz, the bispectrum is flat, as it is dominated by the radio contribution. As frequency increases, the IR contribution becomes increasingly important, first at low multipoles but it finally dominates the whole multipole range at 350 GHz. The *Planck* central CMB channels correspond nearly to the 148 & 219 GHz frequencies. At 148 GHz, we see that the IR contribution is subdominant at 148 GHz except at very low multipoles ($\ell \lesssim 100$). At 219 GHz the IR contribution is important at low multipoles but becomes subdominant beyond $\ell \sim 500$.

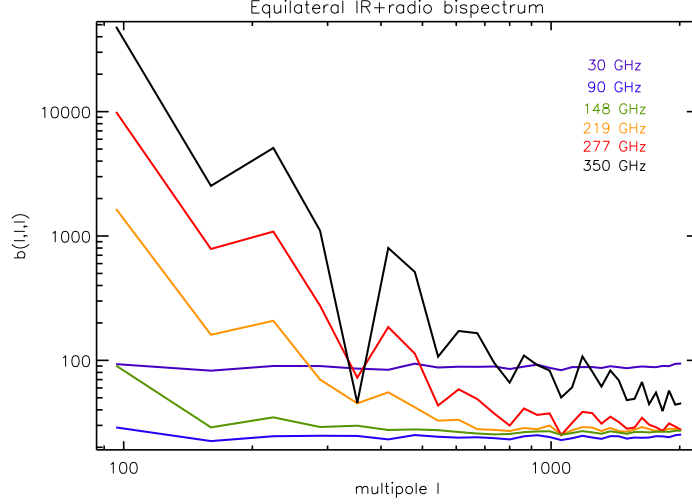


Figure 4.7: IR + radio equilateral bispectrum from Lacasa et al. (2012) with simulations by Sehgal et al. (2010), depending on frequency.

I have compared these measured bispectra with the prediction of the prescription (for radio, IR, IR+radio), where I used the power spectrum measured on the simulations and I fitted for the parameter α . I also computed an expected value of α , base on the sources catalog provided by Sehgal et al. (2010) and on Eq.4.16. I found that the fitted α was consistent with the expected value, although somewhat lower : e.g. at 277 GHz I found $\alpha_{\text{fit}} = 3.1 \cdot 10^{-3}$ and $\alpha_{\text{exp}} = 4.7 \cdot 10^{-3}$. For the superposition of IR and radio signals, Fig.4.8 shows the measured bispectrum at 350 GHz in some configurations (black line), together with the prescription and its cosmic variance (red with error bars). The cosmic variance was computed with the $C_{2 \times 2 \times 2}$ and $C_{3 \times 3}$ terms (see Sect.2.2.5.2), and an $f_{\text{SKY}} = 1/8$ approximation to account for the octant replication in the Sehgal et al. (2010) simulations. Figure 4.8 also shows in blue dashed line what the prescription would yield if the radio and IR populations were considered as 100% correlated random fields.

We see that the prescription reproduces fairly well the measured bispectrum over the whole multipole range and for all configurations. The relative discrepancy between the prescription and the measurement is typically within $\pm 2\%$ at 30 & 90 GHz and increases up to $\pm 30\%$ at 350 GHz. Figure 4.8 shows in fact this latter worst case. We see that the difference between the prescription and the measurement is compatible with being caused by the cosmic variance of the measurement.

Clearly also, the blue dashed curve overestimates the IR+radio bispectrum at high multipoles, while the red curve does not underestimate it. Hence, the IR and radio populations have to be considered independent, as the prescription does. Indeed, the main contribution to the signals does not come from galaxies in the same redshift range, as the radio sources reside at $z \lesssim 1$ while the IR galaxies reside at $z \sim 1 - 5$.

While the radio bispectrum is constant, it is interesting to explore the configuration dependence of the IR bispectrum, which is not readily seen with the previous plots. For this, I have plotted in Fig.4.9 the IR bispectrum at 148 GHz with the parametrisation described in Sect.2.2.4.6.

We remind the reader that each subplot corresponds to a slice in a perimeter bin ($\ell_1 + \ell_2 + \ell_3 \in [P, P + \Delta P]$), with perimeter increasing from left to right and top to

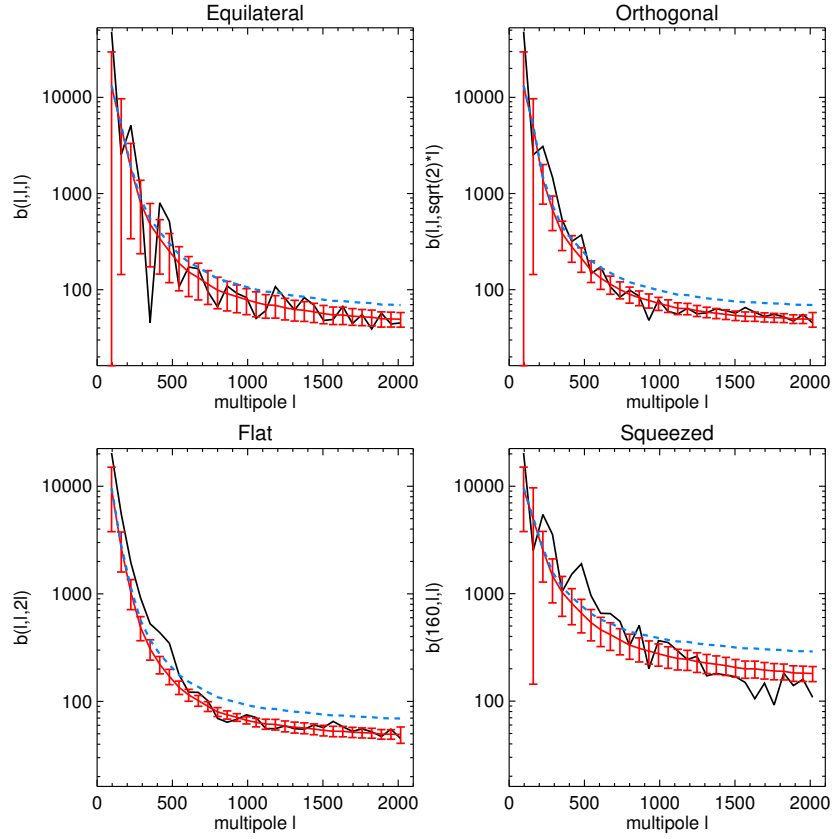


Figure 4.8: From Lacasa et al. (2012), IR + radio bispectrum at 350 GHz in some chosen configurations.

bottom. The color code ranges between violet-blue for the lowest values to red for the highest values. Figure 4.9 shows that the IR bispectrum generally decreases with

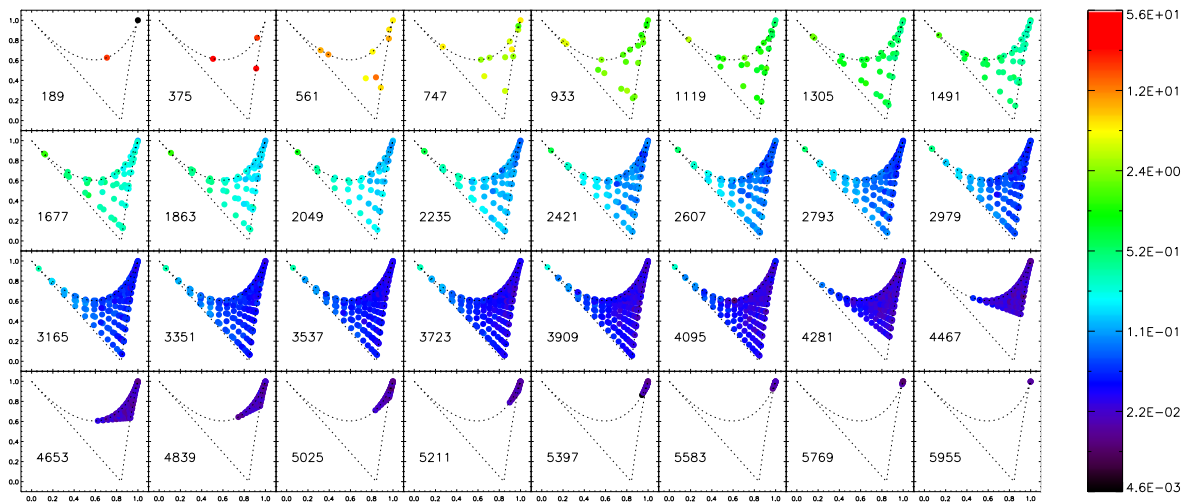


Figure 4.9: From Lacasa et al. (2012), IR bispectrum at 148 GHz in Jy^3/sr plotted with the parametrisation described in Sect.2.2.4.6.

scale, as already noted. Most interestingly in a given perimeter bin, we see that the bispectrum peaks in the squeezed configuration (upper left corner of each subplot). This is an important result, as the CMB bispectrum also peaks in squeezed for local type non-Gaussianity.

Hence, we expect the IR bispectrum to have a shape more correlated⁶ with the CMB one than the radio bispectrum does. This correlation, how it calls for a joint estimation of sources of non-Gaussianity and how it affects the estimation of CMB NG, will be presented in Sect.5.2.1 & 5.2.2.

4.3.3 CIB description with the halo model

The prescription is phenomenological and provides a quick and efficient prediction of the CIB bispectrum. However, a physical modelisation with the halo model would provide an actual model which consistently predicts the power spectrum and bispectrum (and possibly higher order), and hence can be constrained from data.

Several modelisations of the CIB clustering with the halo model have been proposed, all of which focus on the power spectrum only. This section, based on two articles (Lacasa et al. 2013; Pénin et al. 2013, to be submitted), describes the work I have done on modeling CIB polyspectra. We obtain a theoretical description at all orders, and implement it numerically for the bispectrum. This section also shows the improvement in parameter constraint that the bispectrum yield over a power spectrum analysis.

The CIB intensity in a given direction \hat{n} is the line-of-sight integral of the emissivity per comoving volume j :

$$I(\hat{n}, \nu) = \int dr a(z) j(r(z)\hat{n}, z, \nu) \quad (4.18)$$

with j in Jy/Mpc so that $I(\hat{n}, \nu)$ has units of Jy/sr and may be converted to a temperature elevation at CMB frequencies through Planck's law. Through Sect.2.4.4, the angular CIB polyspectra can then be computed from the 3D j polyspectra, as long as the latter are diagonal :

$$\mathcal{P}_{\text{CIB}}^{(n)}(\ell_{1\dots n}) = \int \frac{r^2 dr}{r^{2n}} a^n(z) \mathcal{P}_j^{(n)}(k_{1\dots n}^*, z) \quad (4.19)$$

The IR emissivity traces the galaxy distribution but galaxies at the same redshift or in the same halo may have different IR fluxes. The simplest approach is to consider that the IR flux is stochastic with a distribution given by the number counts. Furthermore, the flux stochasticity is considered independent of the position stochasticity, e.g. the fluxes of two galaxies are not correlated whether they are close together or not. However, the flux distribution depends on the redshift, as the number counts do.

With these assumptions, all terms of the n -th order j polyspectrum which involve n *different* galaxies (i.e. non shot-noise) take the form :

$$\mathcal{P}_j^{(n)}(k_{1\dots n}^*, z) = \bar{j}(z)^n \mathcal{P}_{\text{gal}}^{(n)}(k_{1\dots n}^*, z) \quad (4.20)$$

⁶Note that the correlation notion here is that of bispectrum shapes : whether the shapes have a significant overlap or not. This correlation will be quantified in Sect.5.2.1. This correlation is independent of whether or not the signals are *physically* correlated, and in the IR radio and CMB cases the signals can even be considered independent.

where $\bar{j}(z)$ is the mean IR emissivity, given in terms of the number counts :

$$\bar{j}(z) = \frac{1+z}{\frac{dr}{dz}} \int S \frac{d^2n}{dS dz} dS \quad (4.21)$$

and where the galaxy polyspectrum terms are the non-shot noise ones and can be computed with the diagrammatic method described in Sect.3.4.4.

I have shown in Lacasa et al. (2013) that the shot-noise contributions to $\mathcal{P}_j^{(n)}$ can be tackled with equations to Eq.4.20. For example, if we consider a diagram with a contraction of second order, i.e. the same galaxy is 'hit' twice by the correlation function, we need to consider the average flux squared $\langle S^2 \rangle$ instead of the average flux $\langle S \rangle$ as $\bar{j}(z)$ does. This can be done by defining the p-th order emissivities :

$$j^{(p)}(z) = \frac{(1+z)^p r(z)^{2p-2}}{\frac{dr}{dz}} \int S^p \frac{d^2n}{dS dz} dS \quad (4.22)$$

These are generalisations of the mean emissivity, as e.g., $j^{(1)}(z) = \bar{j}(z)$.

Then, the shot-noise 3D polyspectra should be multiplied by factors $\bar{n}_{\text{gal}}(z)$ in accordance with the degree of p-th order emissivities introduced. For the above example of diagrams with a contraction of second order, we get :

$$\begin{aligned} \mathcal{P}_{j,\text{contrac2}}^{(n)}(k_{1\dots n}^*, z) &= \bar{j}(z)^{n-2} j^{(2)}(z) \bar{n}_{\text{gal}}(z) \mathcal{P}_{\text{gal,contrac2}}^{(n)}(k_{1\dots n}^*, z) \\ &= \bar{j}(z)^{n-2} j^{(2)}(z) \left(\mathcal{P}_{\text{gal,non-shot}}^{(n-1)}(|\mathbf{k}_1 + \mathbf{k}_2|^*, k_{3\dots n}^*, z) + \text{perm.} \right) \end{aligned} \quad (4.23)$$

where I used the resummation of shot-noise diagrams, described in Sect.3.4.4, for the second equality.

Let us note in passing, that this derivation of the shot-noise terms is consistent with the terms which were already known in the literature, i.e. the case where the correlation function 'hits' a single galaxy everytime (while this formalism also allows the derivation of the more complex shot-noise terms which arise at high orders). Indeed e.g. at the power spectrum level we have :

$$C_\ell^{\text{shot}} = \int \frac{dr}{r^2} a^2(z) j^{(2)}(z) \times 1 = \dots = \int S^2 \frac{dn}{dS} dS \quad (4.24)$$

i.e. the well-known shot-noise equation (e.g. Knox et al. 2001).

Now, we can tackle the computation of the CIB bispectrum. All the bispectrum diagrams are shown in Fig.4.10 with, from left to right and top to bottom : 3-halo (3h), 2-halo 3-galaxies (2h3g), 2-halo 2-galaxies (2h2g), 1-halo 3-galaxies (1h3g), 1-halo 2-galaxies (1h2g) and 1-halo 1-galaxy (1h1g).

The third and fifth diagrams (2h2g & 1h2g) are shot-noise diagrams which can be resummed together, and in the following I will call shot2g (for shot-noise with two galaxies) this combination. I will also call simply 2-halo the 2h3g term, 1-halo the 1h3g term, and shot1g (shot-noise with one galaxy) the 1h1g term. The equations for these terms

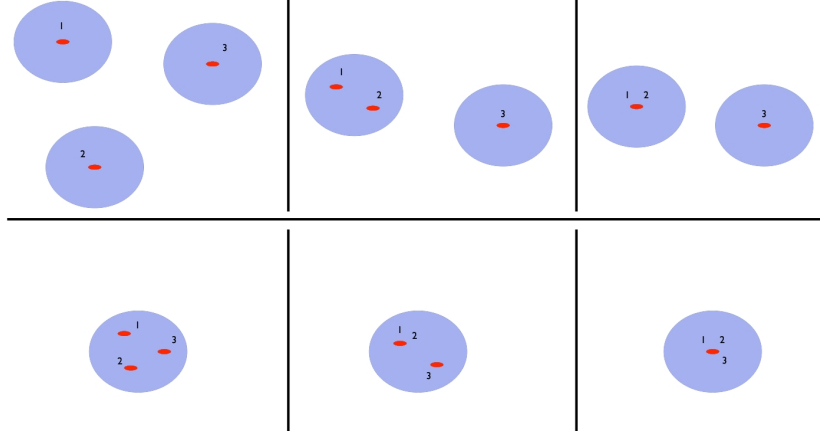


Figure 4.10: Bispectrum diagrams. From left to right and top to bottom : 3-halo (3h), 2-halo 3-galaxies (2h3g), 2-halo 2-galaxies (2h2g), 1-halo 3-galaxies (1h3g), 1-halo 2-galaxies (1h2g) and 1-halo 1-galaxy (1h1g).

are respectively :

$$B_j^{1h}(k_{123}, z) = j^{(1)}(z)^3 \int dM \frac{\langle N_{\text{gal}}(N_{\text{gal}} - 1)(N_{\text{gal}} - 2) \rangle}{\bar{n}_{\text{gal}}^3(z)} \frac{dn_h}{dM} \times u(k_1|M, z) u(k_2|M, z) u(k_3|M, z) \quad (4.25)$$

$$B_j^{2h}(k_{123}, z) = j^{(1)}(z)^3 \int dM_{ab} \frac{\langle N(N-1) \rangle (M_a)}{\bar{n}_{\text{gal}}^2(z)} \frac{\langle N \rangle (M_b)}{\bar{n}_{\text{gal}}(z)} \frac{dn_h}{dM} \Big|_{M_a} \frac{dn_h}{dM} \Big|_{M_b} \times u(k_1|M_a, z) u(k_2|M_a, z) u(k_3|M_b, z) P_{\text{halo}}(k_3, z) + \text{perm.} \quad (4.26)$$

$$B_j^{3h}(k_{123}, z) = j^{(1)}(z)^3 \int dM_{123} \left[\frac{\langle N_{\text{gal}} \rangle (M_i)}{\bar{n}_{\text{gal}}(z)} \frac{dn_h}{dM} \Big|_{M_i} u(k_i|M_i, z) \right]_{i=123} B_{\text{halo}}(k_{123}, z) \quad (4.27)$$

$$B_j^{\text{shot2g}}(k_{123}, z) = j^{(1)}(z)^3 j^{(2)}(z) (P_{\text{gal}}^{\text{non-shot}}(k_1) + P_{\text{gal}}^{\text{non-shot}}(k_2) + P_{\text{gal}}^{\text{non-shot}}(k_3)) \quad (4.28)$$

$$B_j^{\text{shot1g}}(k_{123}, z) = j^{(3)}(z) \quad (4.29)$$

The halo polyspectra are biased versions of the dark matter polyspectra with the local bias scheme described in Sect.3.4.2, while the dark matter polyspectra are computed from perturbation theory. I have been working at tree-level, that is to take the lowest order in the biasing scheme and perturbation theory which gives a non-zero result. For example, for the halo power spectrum we just need to go at first order in local biasing and in perturbation theory (i.e., linear theory described in Sect.3.4.1), we get :

$$P_{\text{halo}}(k|M_1, M_2, z) = b_1(M_1) b_1(M_2) P_{\text{lin}}(k, z) \quad (4.30)$$

For the halo bispectrum, linear theory predicts a vanishing dark matter bispectrum for Gaussian initial conditions. Hence, we need to go to second order in local biasing and/or perturbation theory, we get :

$$B_{\text{halo}}(k_{123}|M_{123}, z) = b_1(M_1) b_1(M_2) b_1(M_3) B_{2\text{PT}}(k_{123}, z) + b_1(M_1) b_1(M_2) b_2(M_3) P_{\text{lin}}(k_1, z) P_{\text{lin}}(k_2, z) + \text{perm.} \quad (4.31)$$

where the dark matter bispectrum at second order in perturbation theory (2PT) is (Fry 1984) :

$$B_{2\text{PT}}(k_{123}, z) = F^s(\mathbf{k}_1, \mathbf{k}_2) P_{\text{lin}}(k_1) P_{\text{lin}}(k_2) + 2 \text{ perm.} \quad (4.32)$$

with the 2PT kernel :

$$F^s(\mathbf{k}_i, \mathbf{k}_j) = \frac{5}{7} + \frac{1}{2} \cos(\theta_{ij}) \left(\frac{k_i}{k_j} + \frac{k_j}{k_i} \right) + \frac{2}{7} \cos^2(\theta_{ij}) \quad (4.33)$$

Note that although Eq.4.33 is singular in the squeezed limit, Eq.4.32 is not, as divergences cancel themselves in the sum over permutations.

Hence, there are two sources for the halo non-Gaussianity : the non-linearity of halo biasing with respect to dark matter, and the non-linearity of gravitational infall producing dark matter NG.

Note that considering primordial NG (PNG) would be more complex as it introduces a momentum dependence of the bias factors (Dalal et al. 2008), we however checked that the corresponding 3-halo term was negligible for f_{NL} up to 50. Furthermore *Planck* results constrained PNG to be consistent with zero.

With these equations, the 2-halo and 3-halo term can be rewritten :

$$B_{\text{gal}}^{2\text{h}}(k_{123}, z) = \mathcal{G}_1(k_1, k_2, z) P_{\text{lin}}(k_3, z) \mathcal{F}_1(k_3, z) + \mathcal{G}_1(k_1, k_3, z) P_{\text{lin}}(k_2, z) \mathcal{F}_1(k_2, z) + \mathcal{G}_1(k_2, k_3, z) P_{\text{lin}}(k_1, z) \mathcal{F}_1(k_1, z) \quad (4.34)$$

$$B_{\text{gal}}^{3\text{h}}(k_{123}, z) = \mathcal{F}_1(k_1, z) \mathcal{F}_1(k_2, z) \mathcal{F}_1(k_3, z) \times [F^s(\mathbf{k}_1, \mathbf{k}_2) P_{\text{lin}}(k_1, z) P_{\text{lin}}(k_2, z) + \text{perm.}] + \mathcal{F}_1(k_1, z) \mathcal{F}_1(k_2, z) \mathcal{F}_2(k_3, z) \times P_{\text{lin}}(k_1, z) P_{\text{lin}}(k_2, z) + \mathcal{F}_1(k_1, z) \mathcal{F}_2(k_2, z) \mathcal{F}_1(k_3, z) \times P_{\text{lin}}(k_1, z) P_{\text{lin}}(k_3, z) + \mathcal{F}_2(k_1, z) \mathcal{F}_1(k_2, z) \mathcal{F}_1(k_3, z) \times P_{\text{lin}}(k_2, z) P_{\text{lin}}(k_3, z) \quad (4.35)$$

where I introduced the notations :

$$\mathcal{F}_i(k, z) = \int dM \frac{\langle N_{\text{gal}}(M) \rangle}{\bar{n}_{\text{gal}}(z)} \frac{dN_h}{dM}(M, z) b_i(M, z) |u(k|M, z)| \quad (4.36)$$

and :

$$\mathcal{G}_1(k_1, k_2, z) = \int dM \frac{\langle N_{\text{gal}}(N_{\text{gal}} - 1) \rangle}{\bar{n}_{\text{gal}}(z)^2} \frac{dN_h}{dM}(M, z) b_1(M, z) |u(k_1|M, z) u(k_2|M, z)| \quad (4.37)$$

In the following, I will call “3hcos” the part of the 3h term containing the 2PT kernel F^s , and simply “3h” the part with \mathcal{F}_2 factors coming from second order halo biasing.

Based on the formalism described above, I have developed a fast and complete a code to compute the galaxy 3D power spectrum and bispectrum, and the CIB angular power spectrum and bispectrum. As stated previously, I use the Sheth & Tormen (1999) mass function and associated bias parameters⁷, the Navarro et al. (1997) halo profile. I use the WMAP7 cosmological parameters (Komatsu et al. 2011). The IR emissivities are computed with number counts by Béthermin & Dole (2011), and the HOD is described

⁷Indeed, this is the most recent mass function for which the second order bias is available.

in Sect.3.4.3. I use fiducial HOD parameters fitted to reproduce the *Planck* CIB angular power spectrum (Planck-Collaboration 2011) : $M_{\min} = 10^{11.5} M_{\odot}$, $M_{\text{sat}} = 10 M_{\min}$, $\sigma_{\log M} = 0.65$ and $\alpha_{\text{sat}} = 1.4$.

Figure 4.11 shows the resulting galaxy bispectrum and its different components in the equilateral configuration at $z=0.1$ and $z=1$. Note that the k -range depends on redshift. Indeed, in view of the computation of the angular bispectrum and to speed up an otherwise time expensive code, we only compute in the code the 3D power spectra and bispectra for modes which project on the multipoles grid (through $k^* = (\ell + 1/2)/r(z)$).

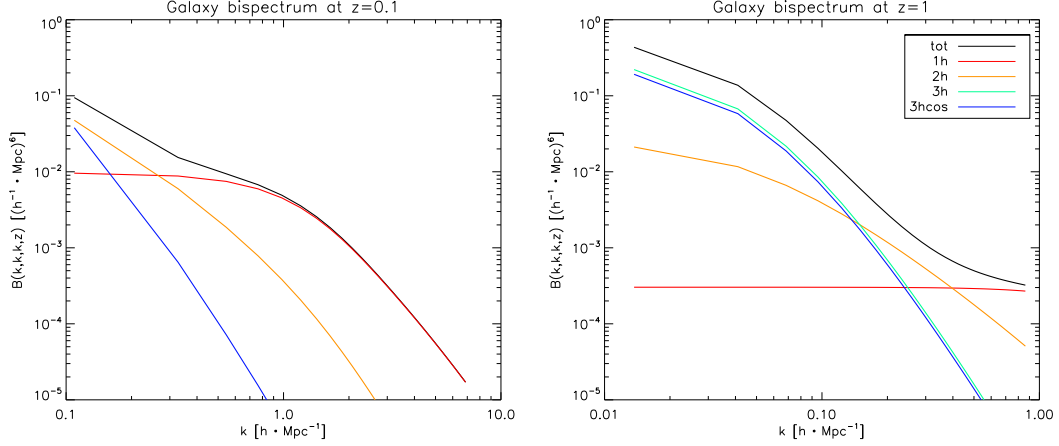


Figure 4.11: From Lacasa et al. (2013), contributions to the galaxy equilateral bispectrum at $z=0.1$ (left panel) and $z=1$ (right panel). Note that the k -range depends on redshift, as these are the modes that will project onto the multipoles of interest (see text for details). The 3h term is negative at $z=0.1$ and thus absent of the left panel.

We see that the 1-halo term is flat on large scales and decreases after $\sim 1 \text{ h} \cdot \text{Mpc}^{-1}$, a scale corresponding to the typical size of a massive halo ($M = 10^{12-13} M_{\odot}$). The 2-halo and 3-halo terms decrease monotonically with scale, with the 3h and 3hcos terms having the steeper slopes while the 2h term has an intermediate slope. Note that the 3h term is negative at low redshifts as, as shown in Fig.3.5, the second order bias is negative on a broader mass range at lower redshifts.

Figure 4.12 shows the evolution of the total galaxy bispectrum in the equilateral configuration, on the range of redshift where there is significant IR emission.

We see that the bispectrum increases considerably with time (whereas the power spectrum increases much more slowly), as non-Gaussianity builds up with non-linear physics entering in action. The figure also shows clearly the projection effect through which the k -range evolves with redshift (particularly at low redshift, afterwards the co-moving distance increases more slowly). Note that the “bump” visible on large scales corresponds to the wavevector k_{eq} ⁸ where the linear power spectrum slope breaks from k to k^{-3} (see Eisenstein & Hu 1999, whose fit I use)⁹.

⁸ k_{eq} is the wavevector whose physical wavelength becomes equal to the Hubble distance at the matter-radiation equality, i.e. $\lambda_{\text{eq}} = c/H(a_{\text{eq}})$.

⁹The Eisenstein & Hu (1999) fit I use does not incorporate the effect of Baryonic Acoustic Oscillations, however these oscillations would be washed out by the redshift integration.

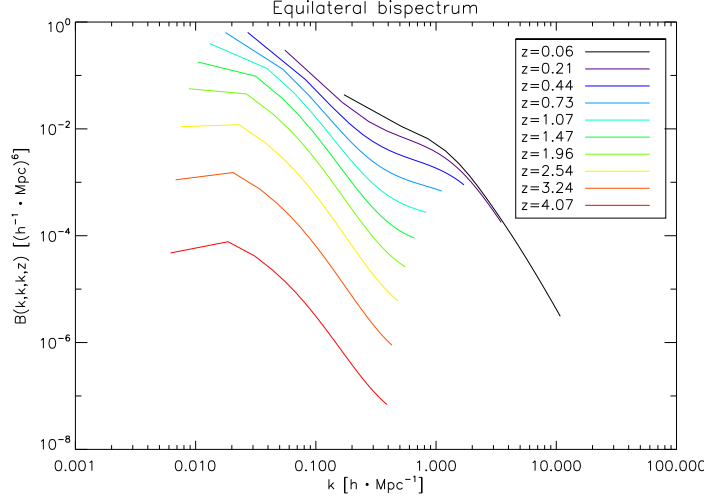


Figure 4.12: From Lacasa et al. (2013), total galaxy equilateral bispectrum as a function of redshift.

Using the IR emissivities computed from number counts by Béthermin & Dole (2011), we can integrate over redshift Eq.4.19 to compute the CIB angular bispectrum. However, we are interested in unresolved galaxies hence we should account for the flux cut effect. Indeed, the flux cut discards the brightest sources, hence it has an effect on the average emissivity and higher order emissivities through Eq.4.22, where the integral runs from $S = 0$ to S_{cut} . As it discards sources, the flux cut also has an effect on the galaxy clustering : there are less galaxies to consider, particularly at low redshift, as they are resolved (and hence masked as we focus on unresolved galaxies). The flux cut effect is modeled with an effective redshift cut in Eq.4.19, considering that galaxies below z_{cut} are resolved, where we have :

$$S_{\text{cut}} = \frac{L_*}{4\pi d_L^2(z_{\text{cut}})} \quad (4.38)$$

where d_L is the luminosity distance and L_* is the typical galaxy IR luminosity (technically : the knee of the luminosity function).

Thus in Eq.4.19 the redshift integral runs from z_{cut} to z_{max} ($z_{\text{max}} = 7$ with the number counts we use, although the contribution of the last redshift bins is negligible).

The CIB total angular bispectrum and its different terms, computed over a range of multipoles of interest for *Planck* ($\ell = 32 \cdots 2048$ with a step $\Delta\ell = 64$), are shown in Fig.4.13 in some chosen configurations at 857 GHz.

We see that the 3h and 3hcos terms (green curves) dominate on large angular scales, the 2-halo term (orange) is important on intermediate scales (towards $\ell \sim 300$ but on a limited multipole range), and the 1-halo term (red) dominates on smaller scales. Finally, the shot-noise term shot1g (dark blue) dominates on the smallest scales, which are not accessible to the *Planck* resolution but are of interest for the *Herschel* telescope.

It is worth emphasizing that these trends do depend on the HOD parameters used. For example, by increasing α_{sat} ¹⁰ we give more weight to massive halos at low redshift. It is quite possible for the 1-halo term to dominate the whole multipole range (as it happens for the tSZ signal at the power spectrum level).

¹⁰to values which are nevertheless not consistent with N-body simulations.

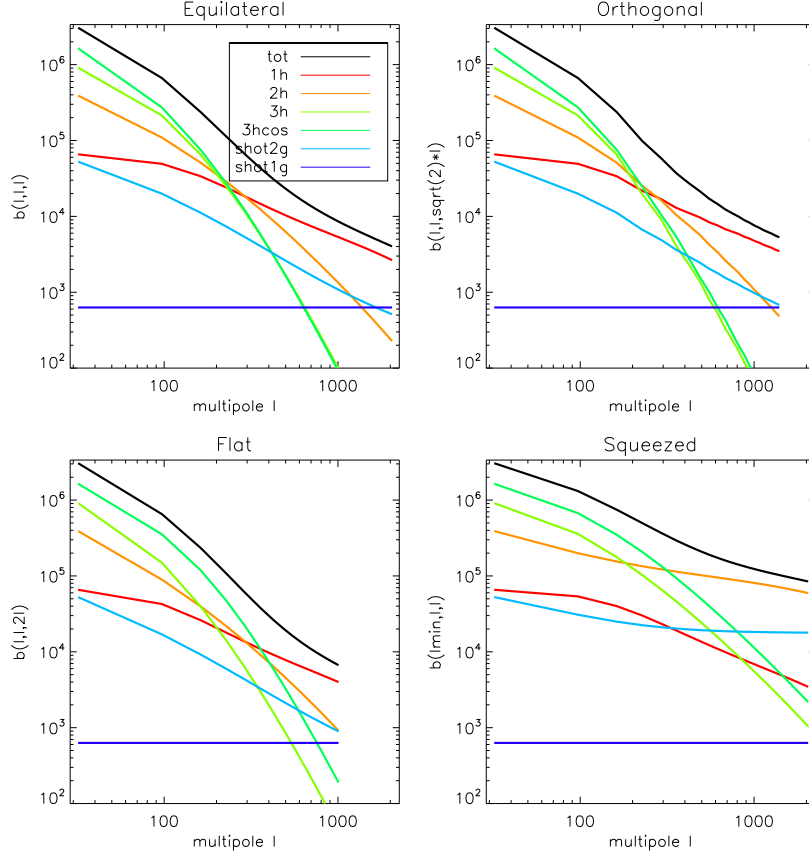


Figure 4.13: From Lacasa et al. (2013), CIB angular bispectrum at 857 GHz in some chosen configurations, and the different terms contributing to it. The black line is the total bispectrum while the color lines give the different terms : shot-noise with 1 galaxy in violet, shot-noise with 2 galaxies in blue, 1-halo in red, 2-halo in orange, 3-halo from non-linear halo biasing in light green, and 3-halo from perturbation theory in darker green.

Interestingly, in the present case we see that the 3hcos term dominates the 3h term in the flat isosceles and squeezed configurations. This is due to the F^s kernel which is more important in flat triangles than in general ones. We also observe that the 2-halo term is often dominated by other terms in most configurations, but it takes its revenge in the squeezed limit, where it dominates at small angular scales ($\ell > 400$). At even smaller angular scales (not shown on the figure), the 2-halo term eventually crosses the shot2g term so that the later dominates the bispectrum in this case.

We can also examine the configuration dependence of the CIB bispectrum and its different terms. In agreement with Sect.4.3.2, we found that the CIB bispectrum peaks in the squeezed configurations. Figure 4.14 shows some of the terms plotted in the parametrisation, note that the color code is logarithmic and adapted to each term.

The upper panel shows the 1-halo term. We see that it is strongly scale dependent but has little dependence on the configuration, as it is nearly constant in a perimeter bin. The middle panel shows the 3hcos term. It also has a strong scale dependence and

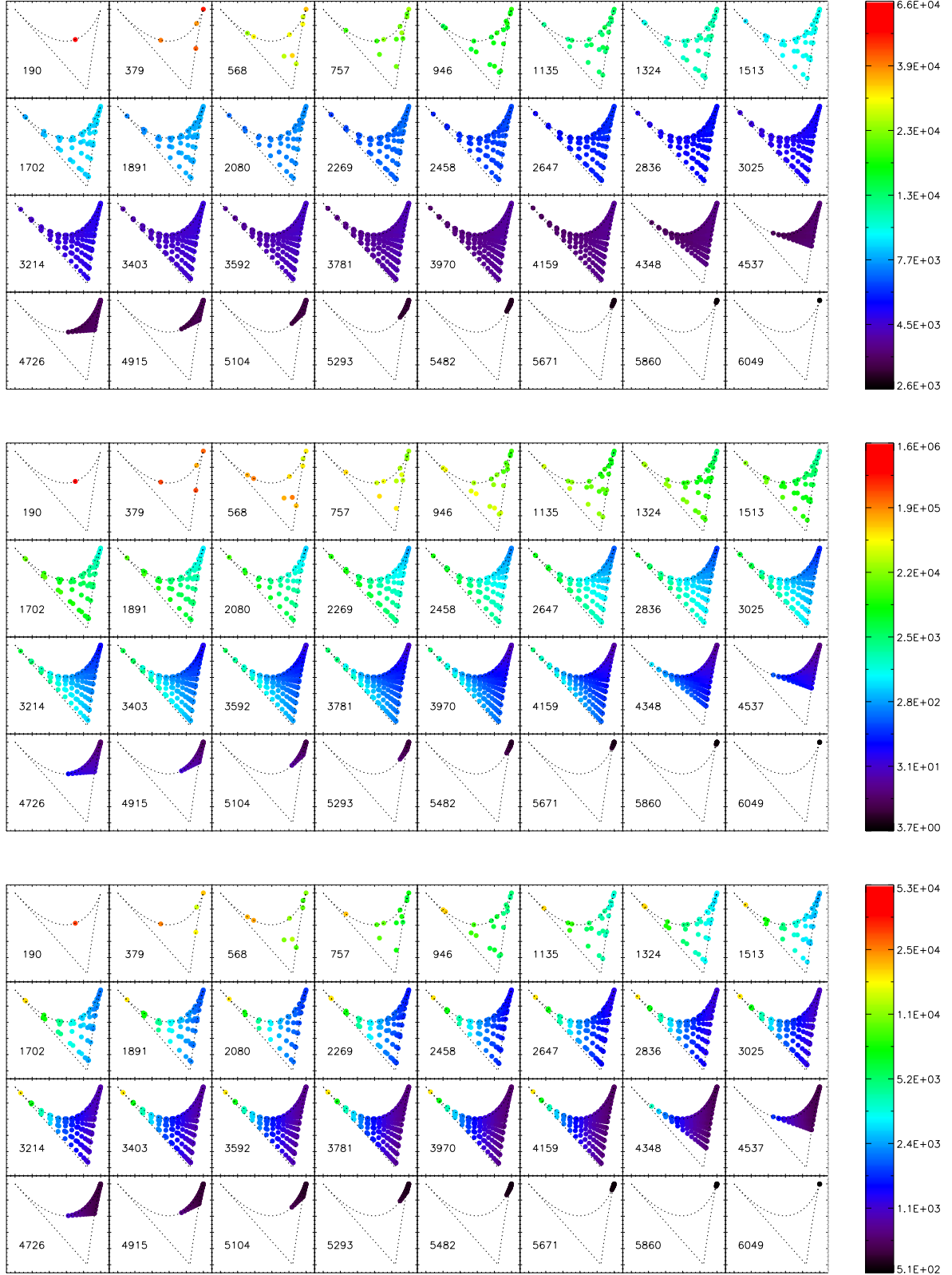


Figure 4.14: From Lacasa et al. (2013), some terms of the CIB angular bispectrum plotted in the parametrisation. From top to bottom : 1h, 3hcos, shot2g. The color code is logarithmic and adapted to each term. Each of the 32 subplots corresponds to a perimeter bin, with the center of the bin indicated in its lower left corner ; the squeezed triangles are present in the upper left corner, the equilateral triangles in the upper right and the flat isosceles in the lower right. 88

shows a peculiar configuration dependence : it peaks in the flat triangles (from squeezed to flat isosceles), because of the F^s kernel, but the squeezed configuration takes over the flat isosceles in high perimeter bins, because of the halo profile which decreases with ℓ . The lower panel shows the shot2g term. It peaks strongly in the squeezed configuration. The 2-halo term shows a similar behaviour, although with a different dynamic range and less difference between the squeezed and the other configurations. The 3-halo term also peaks in squeezed but less strongly than the 2-halo and the shot2g terms. The last term is the shot1g one but its configuration dependence is trivial as it is constant. Hence, we see that the scale and configuration dependence of each of the CIB bispectrum terms are different ; thus the different terms may be discriminated in measurements using this information.

In Pénin et al. (2013), we have looked at the influence of the IR emissivity on the bispectrum, using three different IR emissivity models. We have shown that the IR emissivity has few influence on the configuration dependence of the bispectrum terms, however it strongly affects the scale dependence and the amplitude of the bispectrum. In particular we have shown that models which may be degenerate at the power spectrum level (i.e. producing indistinguishable C_ℓ once the HOD is fitted) may be distinguished at the bispectrum level, as the bispectra they predict may differ by up to a factor of 10.

I have also compared in Lacasa et al. (2013) the CIB bispectrum computed from the halo model with the one derived from the empirical prescription. For this purpose, I used the same amplitude parameter α as the one deduced from Sehgal et al. (2010) simulations (see Sect.4.3.2). Furthermore I used the power spectrum computed with the halo model. In doing so, I am testing the functional form of the prescription. The comparison is displayed in Fig.4.15.

The prescription reproduces the overall shape and the order of magnitude of the halo-model bispectrum. The prescription has generally a higher amplitude than the halo-model bispectrum, which can be compensated by rescaling the amplitude parameter α by a factor $\sim 3/4$. Interestingly, the halo-model bispectrum departs from the prescription at multipoles $\ell > 300$ in the squeezed limit, that is when the 2-halo term starts to dominate. Hence, the prescription is a good phenomenological first order approach, but as expected a comprehensive modeling of the bispectrum based on the halo model is necessary for detailed studies and to analyse actual measurements.

Indeed, the goal of using a physically-base model of the bispectrum is to constrain it with observations. To this end it is interesting to concentrate on the variation of the CIB bispectrum with model parameters. For illustration, Fig.4.16 shows the variation of the CIB bispectrum and its different terms, in the equilateral configuration, when one HOD parameter is varied while the others are fixed.

From left to right we vary respectively : M_{\min} , M_{sat} and α_{sat} . These parameters govern respectively the minimal mass for a halo to contain a (central) galaxy, the minimal mass for a halo to contain satellite galaxies, and the number of satellite at high mass ($M \gg M_{\text{sat}}$). For example, we see that varying α_{sat} has a strong effect on the 1-halo term, and slightly less effect on the other terms. The 2-halo term is the most affected, and the 3-halo term (which here regroups the plain 3h and the 3hcos) is the least affected. Indeed increasing α_{sat} affects mostly the massive halos where the number of galaxies increases strongly ($N_{\text{sat}} \propto M^{\alpha_{\text{sat}}}$). As the 1-halo term is cubic with the number of galaxies

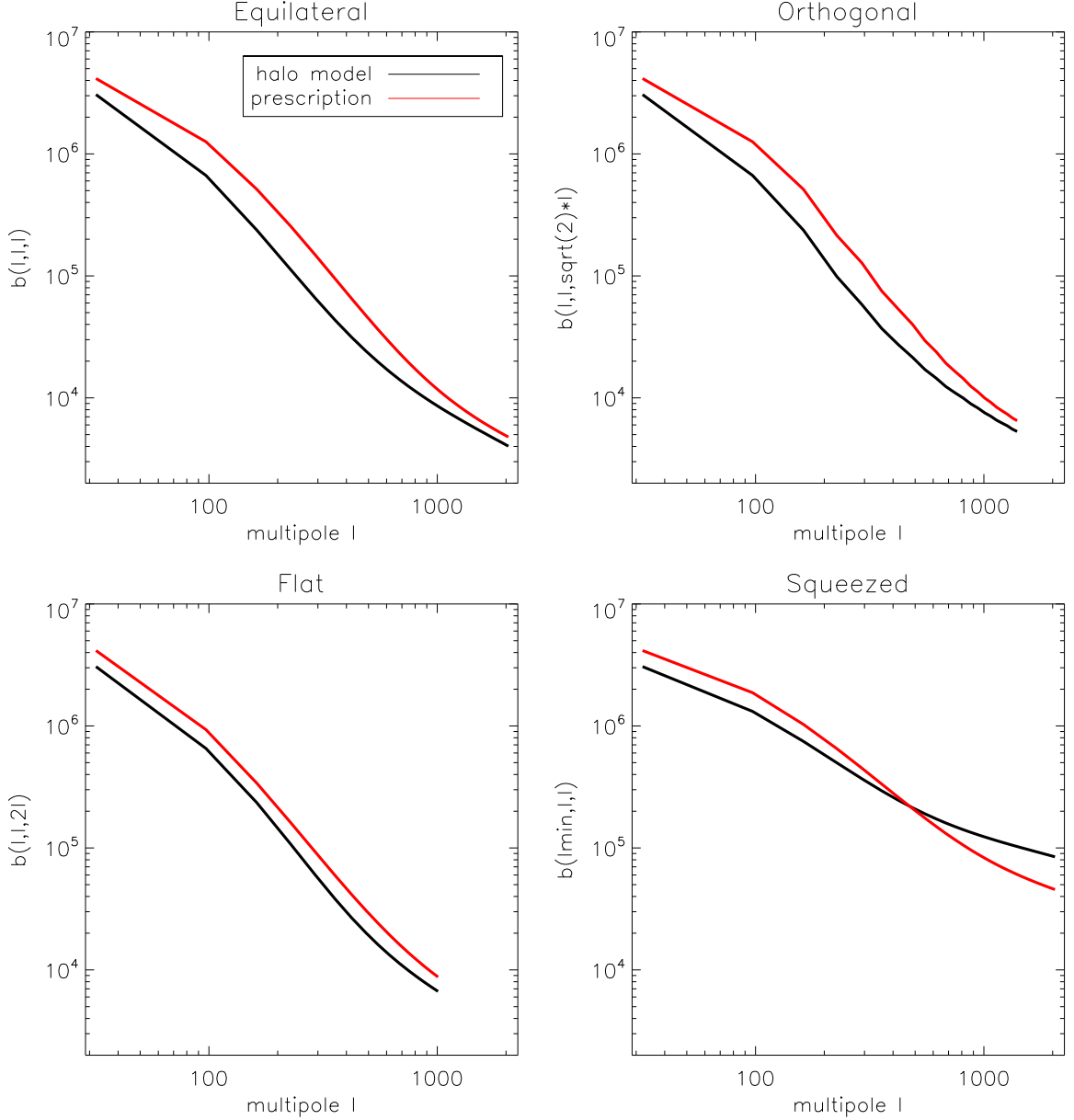


Figure 4.15: From Lacasa et al. (2013), comparison between the halo-model bispectrum and the empirical prescription at 857 GHz in some chosen configurations.

at a given mass ($\langle N(N-1)(N-2) \rangle$ factor), increasing α_{sat} blows up b^{lh} .

We have observed that the variation of HOD parameters has more effect on the bispectrum than on the power spectrum. Hence, the bispectrum seems to be a powerful observable to constrain them. This is to be balanced however by the fact that, from the measurement point of view, bispectrum coefficients have much lower SNR than power spectrum ones, and the bispectrum is more difficult to measure. To quantify the constraint that the power spectrum and bispectrum may respectively yield on the HOD parameters, we have conducted in Pénin et al. (2013) a Fisher analysis. The Fisher

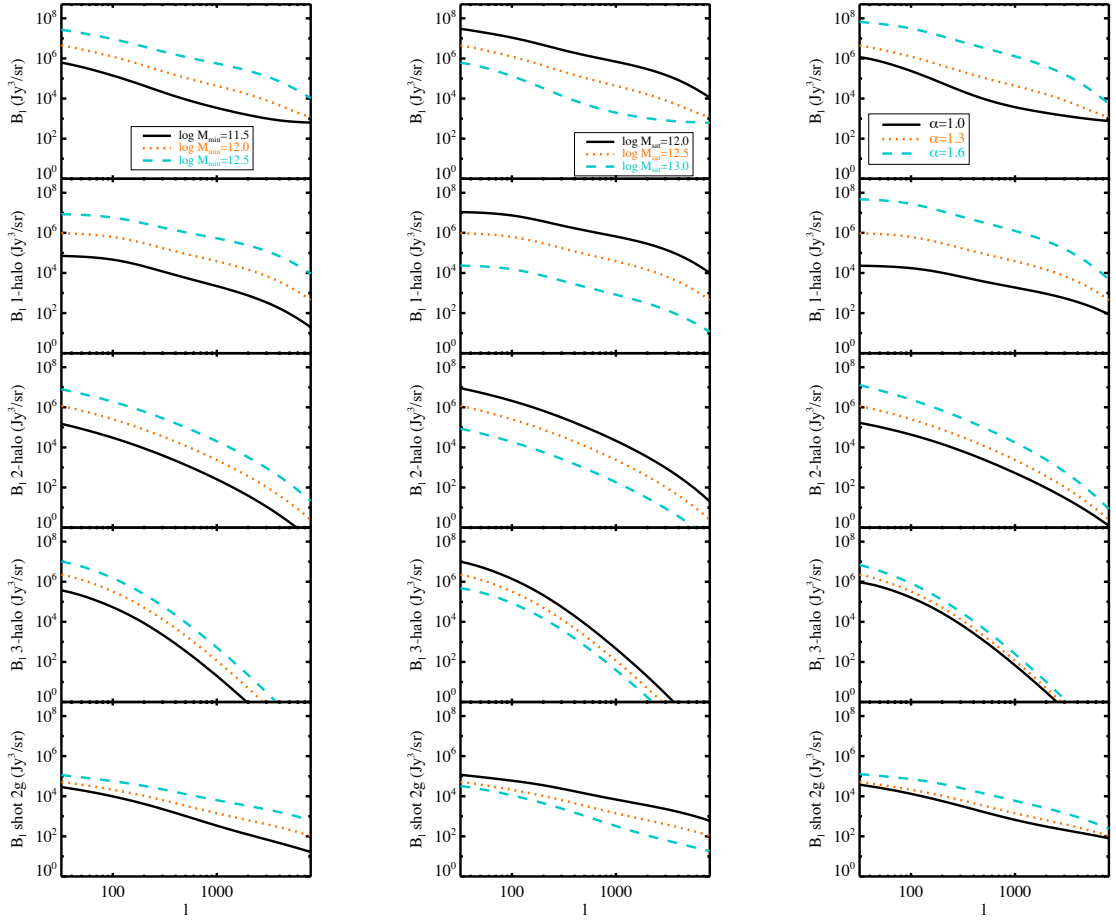


Figure 4.16: From Pénin et al. (2013), variations of the CIB equilateral angular bispectrum and its different terms with the HOD parameters. From left to right : varying M_{\min} , M_{sat} and α_{sat} . The solid line shows the bispectrum for the lowest value of the parameter of interest, the dotted line for an intermediate value, and the dashed line for the highest value of the parameter.

matrix of a set of parameters $(\theta_\alpha)_\alpha$ has entries :

$$F_{\alpha\beta} = \frac{\partial \mathcal{O}}{\partial \theta_\alpha}^T \cdot C^{-1} \cdot \frac{\partial \mathcal{O}}{\partial \theta_\beta} \quad (4.39)$$

where \mathcal{O} is the observable (in our case the power spectrum or the bispectrum) ordered as a vector, and C is its covariance matrix. If the observable can be considered to vary linearly with parameters within the error bars (e.g., if the error bars are sufficiently small), then the likelihood of the parameters $(\theta_\alpha)_\alpha$ is Gaussian and F is its covariance matrix. In other words $F_{\alpha\alpha}^{-1}$ is the error bar on the parameter θ_α when fitting to data. The off-diagonal coefficients give the correlation of error bars on different parameters. To compute the Fisher matrices, we considered binned estimations of the power spectrum and bispectrum with $\ell = 32 \dots 2048$ and $\Delta\ell = 64$, we considered *Planck* Bluebook's beam and noise values (The Planck Collaboration 2006) and $f_{\text{sky}} = 50\%$ ¹¹. For the

¹¹In the f_{SKY} approximation, lower sky fractions increase the error bars by a factor $\sqrt{f_{\text{SKY}}}$ but do

bispectrum we considered the $C_{2 \times 2 \times 2}$ and $C_{3 \times 3}$ terms of the covariance matrix (see Sect.2.2.5).

Then, one can compute from the Fisher matrix the 1, 2 or 3 σ contours whose interior correspond to the 68, 95 or 99 % confidence level (C.L.) respectively. In a 2-dimensional slice of the parameter space, these contours are ellipses. Fig.4.17 shows this ellipses for the three HOD parameters α_{sat} , M_{min} and M_{sat} , at respectively 217 GHz (left panel) and 857 GHz (right panel).

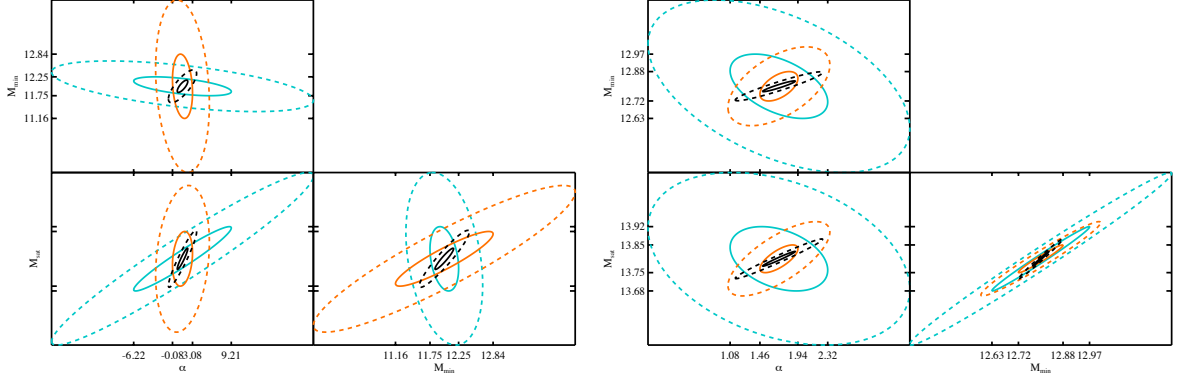


Figure 4.17: Fisher matrix ellipses of HOD parameter constraints, with the power spectrum (blue), the bispectrum (orange) and combined (black). Solid lines are 1σ contours and dashed are 2σ contours. Left figures are at 217 GHz and right at 857 GHz.

We see that the bispectrum brings very interesting constraint on the HOD parameters. At 217 GHz, its degeneracies are complementary to those of the power spectrum, so that the combination of both improves greatly the error bars, in an ideal case without Galactic contamination in particular. Indeed, the error bar on M_{min} is decreased by 30% at 217 GHz and by a factor ~ 5.5 at 857 GHz, by using a combined constraint compared to a power spectrum only constraint. For M_{sat} the corresponding improvement is a factor ~ 3 at 217 GHz and a factor ~ 4 at 857 GHz ; for α_{sat} the improvements are a factor ~ 7.5 at 217 GHz and ~ 3 at 857 GHz. Note also that at 857 GHz, the bispectrum constraint on the parameters are better than those of the power spectrum. However at this frequency the measurement is severely impaired by the contamination by Galactic dust, as will be described in Sect.5.2.3.

The fact that the bispectrum brings more stringent constraint at 857 GHz than at 217 GHz can be understood easily. For the CIB, the signal at higher frequencies has more contribution from low redshift sources (see e.g. Pénin et al. 2012). As the bispectrum is highly sensitive to low redshifts, it is particularly suited for the CIB at 857 GHz.

Thus, we have shown that the CIB bispectrum is potentially a well-suited observable to constrain CIB models. I have built a powerful theoretical framework to predict the CIB clustering at high orders with a consistent and physically motivated model. Furthermore, I have developed a fast and efficient code to compute the CIB angular power spectrum and bispectrum, which can hence be used to constrain the model with measurement of both quantities.

not change the relative power of the power spectrum and bispectrum constraints.

Chapter 5

Measuring non-Gaussianity

Contents

5.1	Observational considerations for <i>Planck</i>	94
5.2	Non-Gaussianity from extragalactic point-sources	95
5.2.1	IR amplitude and joint NG constraints	95
5.2.2	Contamination of primordial NG estimation	101
5.2.3	Measuring the CIB NG with <i>Planck</i> data	105
5.2.3.1	Small fields	106
5.2.3.2	GASS field : bispectrum estimation pipeline	107
5.2.3.3	GASS field : results and post-treatment	111
5.2.3.4	GASS field : exploitation of the measurements	115
5.3	Measuring the thermal Sunyaev-Zel'dovich NG	117
5.3.1	On simulations	117
5.3.2	<i>Planck</i> tSZ data : characteristics and processing pipeline	118
5.3.3	<i>Planck</i> tSZ data : results and characterisation	123

In the previous chapter, I have described extragalactic foregrounds to the CMB, and e.g., I have shown how high order moments contain valuable information to constrain the CIB model. In this chapter, I describe the statistical methods I have proposed and the data analysis I have performed, for the measurement of the non-Gaussianity of extragalactic point-sources and of the thermal SZ signal with *Planck* data.

5.1 Observational considerations for *Planck*

This chapter concerns data analysis, in particular with *Planck* data. We thus need to account for observational limitations which appear in this context. The first limitation is the necessity of a mask due to the astrophysical contamination of areas of the sky, by the Galactic emission or by point-sources. The problem of the inherent incomplete sky coverage was tackled in Sect.2.2.4. There are then two limitations due to the instrumental characteristics that need to be considered :

- Resolution : any instrument has a finite angular resolution, due in particular to the diffraction limit associated to the mirror width. Thus the observation of a point-like signal is not a point-like image but a smoother spot. This image is called the point-spread function or beam. If the beam is independent of the pointing direction, the observed sky is a convolution of the “true sky” by the beam. Furthermore if the beam is rotationally invariant, the observed harmonic coefficients read :

$$a_{\ell m}^{\text{obs}} = a_{\ell m}^{\text{sky}} \times B_{\ell} \quad (5.1)$$

A common and legitimate assumption is that the beam is Gaussian, in which case it is characterised by its full width to half maximum (FWHM) and we have :

$$B_{\ell} = \exp \left(-\ell(\ell+1) \frac{\text{FWHM}^2}{16 \ln(2)} \right) \quad (5.2)$$

with FWHM in rad. FWHM measured for the *Planck* beams are given by Planck Collaboration VII (2013) (they range between 4 and 10 arcmin for the High Frequency Instrument and between 13 and 33 arcmin for the Low Frequency Instrument).

- Noise : even with a perfect acquisition chain, there is a fundamental limit to the detector precision which is photon noise. Photon noise arises from the fact that the number of photons hitting the detector in a given time interval is random (following a Poisson law, see Sect.1.1.4.2). This produces variations of detected power and hence noise in the inferred sky intensity. There are many other possible sources of noise along the acquisition chain, but *Planck* has been designed so that the fundamental photon noise is its main instrumental limitation. If we consider the noise contribution, Eq.5.1 is modified into :

$$a_{\ell m}^{\text{obs}} = a_{\ell m}^{\text{sky}} \times B_{\ell} + a_{\ell m}^{\text{noise}} \quad (5.3)$$

Because the CMB monopole is dominant, we can consider that the average number of photons is isotropic, and hence $a_{\ell m}^{\text{noise}}$ is a random variable independent of $a_{\ell m}^{\text{sky}}$.

Furthermore, this number of photons is large so $a_{\ell m}^{\text{noise}}$ can be considered Gaussian through the central limit theorem. Finally, if all areas of the sky are surveyed during the same time, the noise is isotropic and is hence solely characterised by its angular power spectrum C_{ℓ}^{noise} . This last point is not the case for *Planck* due to its scanning strategy, but it is a common assumption, that we will consider at first order.

In the following, we note with calligraphical sign the power spectrum accounting for the noise and the beam effects :

$$\mathcal{C}_{\ell} = C_{\ell} B_{\ell}^2 + C_{\ell}^{\text{noise}} \quad (5.4)$$

We also note the corresponding quantity in bold for the bispectrum :

$$\mathbf{b}_{123} = b_{\ell_1 \ell_2 \ell_3} B_{\ell_1} B_{\ell_2} B_{\ell_3} \quad (5.5)$$

Let us note that this last equation considers that the noise is Gaussian so that its bispectrum vanishes. I have indeed computed non-Gaussianity estimates on realist noise simulations by the *Planck* collaboration and found them consistent with zero.

Although the bispectrum (and higher orders) estimation is not biased by the noise, contrary to the power spectrum, its error bars are increased due to the noise. For example the $C_{2 \times 2 \times 2}$ term of the bispectrum covariance contains a product of power spectra, and these power spectra include a noise contribution as in Eq.5.4.

Observationally, the noise level can be estimated from the data itself by the so-called jackknife procedure. In the case of *Planck*, the scanning strategy produces data in the form of rings of the sky. Each pixel of the sky is observed several times (on average $\sim 1,000$ times), so maps may be produced from the first half of these rings and compared to the map produced with the other half. These maps are called respectively first and last ring, and their difference (the jackknife) represents instrumental noise as the sky signal is cancelled. I will also make use of these maps for the CIB NG measurement in Sect.5.2.3.

5.2 Non-Gaussianity from extragalactic point-sources

In Sect.4.3.2 I presented the characterisation of the non-Gaussianity from extragalactic point-sources (radio and IR) theoretically and on simulations. The following sections describe the statistical analysis I have performed to measure this non-Gaussianity in the *Planck* context.

5.2.1 IR amplitude and joint NG constraints

As shown in Sect.4.3.2, the CIB bispectrum can be an important contribution to the total bispectrum of the extragalactic signals, especially above 217 GHz. Furthermore, it peaks in the squeezed configurations, i.e., the same as the CMB bispectrum for local type PNG. Hence, the estimation of primordial, CIB and radio non-Gaussianity may interfere with each other, i.e. each separate estimation may be contaminated by the others. For this reason, we need a tool to estimate the amplitude of the CIB bispectrum in a map, and a tool which can be combined consistently with those used to quantify

the primordial and radio NG. This tool also needs to be computationally fast, to enable numerous tests on simulations as needed for *Planck* analysis.

Fast estimators for the parameters f_{NL} and b_{PS} , characterising respectively the amplitude of the primordial and radio bispectra, have been proposed by Komatsu et al. (2005). In Lacasa & Aghanim (2013), I have proposed an estimator for the CIB, following the same principles.

To this end, the parameter A_{IR} is defined as the amplitude of the IR bispectrum compared to a template derived from the empirical prescription Eq.4.15. In practice, for a full-sky map where the CIB is assumed to be the sole NG source, the amplitude of the bispectrum is measured by minimising the χ^2 of the measured bispectrum to the template. This yields :

$$\chi^2(A_{\text{IR}}) = \sum_{\ell_1 \leq \ell_2 \leq \ell_3} \frac{(\mathbf{b}_{123} - A_{\text{IR}} \mathbf{b}_{123}^{\text{IR}})^2}{\sigma^2(\ell_1, \ell_2, \ell_3)} \quad (5.6)$$

with :

$$\sigma^2(\ell_1, \ell_2, \ell_3) = \frac{\mathcal{C}_{\ell_1}^{\text{tot}} \mathcal{C}_{\ell_2}^{\text{tot}} \mathcal{C}_{\ell_3}^{\text{tot}}}{N_{\ell_1 \ell_2 \ell_3}} \times \Delta_{\ell_1 \ell_2 \ell_3} \quad (5.7)$$

being the bispectrum variance in the weak non-Gaussianity limit (the term $C_{2 \times 2 \times 2}$ of the covariance, see Sect.2.2.5.2).

In this expression, $\mathcal{C}_{\ell}^{\text{tot}}$ is the measured power spectrum, and we have :

$$\begin{aligned} \Delta_{\ell_1 \ell_2 \ell_3} &= 1 + \delta_{\ell_1 \ell_2} + \delta_{\ell_1 \ell_3} + \delta_{\ell_2 \ell_3} + 2 \delta_{\ell_1 \ell_2 \ell_3} \\ &= \begin{cases} 6 & \text{for equilateral triangle} \\ 2 & \text{for isosceles triangle} \\ 1 & \text{for general triangle} \end{cases} \end{aligned} \quad (5.8)$$

By minimising the χ^2 , we get the maximum likelihood estimator :

$$\hat{A}_{\text{IR}} = \sum_{\ell_1 \leq \ell_2 \leq \ell_3} \frac{\mathbf{b}_{123} \mathbf{b}_{123}^{\text{IR}}}{\sigma^2(\ell_1, \ell_2, \ell_3)} \times \sigma^2(\hat{A}_{\text{IR}}) \quad (5.9)$$

with :

$$\sigma^2(\hat{A}_{\text{IR}}) = \left(\sum_{\ell_1 \leq \ell_2 \leq \ell_3} \frac{(\mathbf{b}_{123}^{\text{IR}})^2}{\sigma^2(\ell_1, \ell_2, \ell_3)} \right)^{-1} \quad (5.10)$$

being a normalisation factor which gives the error-bar of the A_{IR} estimator.

The bispectrum derived from the prescription, Eq.4.15, is separable. That is, it can be written in the form $b(\ell_1, \ell_2, \ell_3) = \sum_i f_i(\ell_1) g_i(\ell_2) h_i(\ell_3) + \text{perm.}$ This is the case for the prescription bispectrum with the sum reduced to one element and $f_i(\ell) = g_i(\ell) = h_i(\ell) = \alpha^{1/3} \sqrt{C_{\ell}^{\text{IR}}}$. Hence, a simpler computation of the numerator of Eq.5.9 can be devised, inspired by Komatsu et al. (2005). Indeed, defining the filtered map :

$$F(\mathbf{n}) = \sum_{\ell m} \frac{\alpha^{1/3} \sqrt{C_{\ell}^{\text{IR}}} b_{\ell}}{\mathcal{C}_{\ell}^{\text{tot}}} a_{\ell m} Y_{\ell m}(\mathbf{n}) \quad (5.11)$$

and :

$$\mathcal{S}_{\text{IR}} = \int d^2 \mathbf{n} F(\mathbf{n})^3 \quad (5.12)$$

some algebra leads to :

$$\mathcal{S}_{\text{IR}} = \sum_{\ell_1 \ell_2 \ell_3} \frac{\hat{\mathbf{b}}_{123} \mathbf{b}_{123}^{\text{IR}}}{\mathcal{C}_{\ell_1}^{\text{tot}} \mathcal{C}_{\ell_2}^{\text{tot}} \mathcal{C}_{\ell_3}^{\text{tot}}} N_{123} = 6 \times \sum_{\ell_1 \leq \ell_2 \leq \ell_3} \frac{\hat{\mathbf{b}}_{123} \mathbf{b}_{123}^{\text{IR}}}{\sigma^2(\ell_1, \ell_2, \ell_3)} \quad (5.13)$$

The computations of the filtered map and of \mathcal{S}_{IR} take respectively $\mathcal{O}(N_{\text{pix}}^{3/2})$ and $\mathcal{O}(N_{\text{pix}})$ operations, while the denominator of Eq.5.9 is independent of the map and can be pre-computed. Hence, this approach is much faster than a full bispectrum analysis which scales as $\mathcal{O}(N_{\text{pix}}^{5/2})$ operations.

In realistic cases of CMB analysis, the statistical isotropy of the signal is broken by inhomogeneous noise –e.g., due to the scanning strategy– or by masking large areas of the sky –e.g., the galactic plane. In these cases, the estimator of the IR-bispectrum amplitude is no longer optimal and it is biased in a non-trivial way. Nevertheless and similarly to the case of the f_{NL} , the bias and the lack of optimality can both be tackled through adapted modifications to the estimator, similar to the bispectrum case (see Sect.2.2.4.3 & 2.2.4.4).

Bias

When a covariance matrix of the map(s) can be estimated, Creminelli et al. (2006) have shown that Wiener filtering of the map(s) will debias the non-Gaussian estimator from anisotropic contaminants or noise. Specifically, if \mathbf{C} is the estimated covariance matrix in harmonic space, the estimator Eq. 5.9 can be debiased by applying the modification :

$$\frac{a_{\ell m}}{\mathcal{C}_{\ell}^{\text{tot}}} \rightarrow (\mathbf{C}^{-1} \cdot \mathbf{a})_{\ell m} = \sum_{\ell' m'} C_{\ell m, \ell' m'}^{-1} a_{\ell' m'} \quad (5.14)$$

in the filtered map, Eq. 5.11. The denominator of the estimator Eq.5.10 must undergo a consistent modification. This denominator is then most easily computed through simulations. Note that $(\mathbf{C}^{-1} \cdot \mathbf{a})_{\ell m}$ reduces to $\frac{a_{\ell m}}{\mathcal{C}_{\ell}^{\text{tot}}}$ in the isotropic case.

Variance

When isotropy is broken, the 3-point correlation function used to define the bispectrum no longer has minimal variance. It must be replaced by the Wick product of the three harmonic coefficients (Donzelli et al. 2012) :

$$a_1 a_2 a_3 \rightarrow a_1 a_2 a_3 - \langle a_1 a_2 \rangle a_3 - \langle a_1 a_3 \rangle a_2 - \langle a_2 a_3 \rangle a_1 \quad (5.15)$$

These additional terms are the ones which also demand a linear term for the bispectrum estimation, as described in Sect.2.2.4.3. These terms vanish in the isotropic case.

Only the 2-point correlation function is to be considered for the expectation values $\langle \rangle$ in Eq.5.15. Therefore in practice, the latter are obtained from a sufficiently large number of Gaussian realisations with the same power spectrum as that of the signal considered. Inputting the linear corrections given by Eq. 5.15 into the expression of S_{IR} in Eq. 5.13 yields :

$$S_{\text{IR}} \rightarrow \tilde{S}_{\text{IR}} = \int d^2 \mathbf{n} F(\mathbf{n})^3 - 3 \times \int d^2 \mathbf{n} F(\mathbf{n}) G(\mathbf{n}) \quad (5.16)$$

with :

$$G(\mathbf{n}) = \langle F(\mathbf{n})^2 \rangle_{\text{MC}} \quad (5.17)$$

where the brackets $\langle \rangle_{\text{MC}}$ stand for the average over Gaussian simulations.

I now present the expected detection significance of the CIB NG in terms of the signal-to-noise ratio (SNR) for the A_{IR} estimator, through $\text{SNR} = 1/\sigma(\hat{A}_{\text{IR}})$. To this end, I use the values of α and C_{ℓ}^{IR} measured on the simulations of Sehgal et al. (2010), see Sect.4.3.2, with a maximum multipole $\ell_{\text{max}} = 2048$.

First, I present an ideal case : a full-sky CIB map without noise and with perfect angular resolution. In this case, one sees that the signal-to-noise ratio increases with the number of available bispectrum configurations :

$$\text{SNR} \propto \sqrt{N_{\text{tot}}(\ell_{\text{max}})} \quad \text{with} \quad N_{\text{tot}}(\ell_{\text{max}}) = \sum_{\ell_{\text{min}} \leq \ell_1 \leq \ell_2 \leq \ell_3 \leq \ell_{\text{max}}} N_{\ell_1 \ell_2 \ell_3} \quad (5.18)$$

The resulting SNR at the frequencies of the simulations by Sehgal et al. (2010) are shown in the first line of Table 5.1. The obtained SNR are of order ~ 1000 and do not vary much with frequency.

frequency (GHz)	150	220	280	350
Ideal case	1218	1157	1161	1159
Full-sky IR	15.5	98	336	833
Full-sky with CMB	0.39	6.7	55	387
50% sky with CMB	0.28	4.7	39	274
50% sky with 10% CMB	3.67	45.9	210	577

Table 5.1: Expected SNR at $\ell_{\text{max}} = 2048$. Ideal case stands for a full-sky cosmic-variance limited CIB map without noise nor contaminations. All following rows adopt the instrumental characteristics specified in Table 5.2. Full-sky IR represents the case of the CIB map as seen by the instrumental set-up. Full-sky with CMB is when CMB is present, contaminating the CIB map. 50% sky with CMB stands for the case of masking half of the sky. 50% sky with 10% CMB stands for a half sky CIB map where 90% of the CMB (in amplitude) has been removed.

In a more realistic case, the CIB map is convolved by the instrumental beam and the signal is contaminated by noise, but we do not include astrophysical contamination. To model these effects, I used a Gaussian beam and white noise whose values are specified in Table 5.2. Both are representative of a *Planck*-like experiment. I also present partial-sky coverage cases, assuming that the optimisation and the de-biasing described previously have been applied, the SNR scales as $f_{\text{SKY}}^{-1/2}$.

Beam	Noise	C_{ℓ}^{noise}
5 arcmin	$10^{-8} \Delta T/T \cdot \text{sr}^{1/2}$	$7.4 \cdot 10^{-4} \mu\text{K}^2 \cdot \text{sr}$

Table 5.2: Instrumental specifications used throughout this section.

The Table 5.1 shows the case of a convolved but “perfect” full-sky CIB map, a full-sky case but including CMB contamination, a case with a 50% sky-fraction mimicking a galactic mask, and finally the case where 90% of the CMB is removed. As a matter of fact, component separation methods can efficiently estimate the CMB signal from

multifrequency observation with an error in amplitude below 10% (see Delabrouille & Cardoso 2007; Remazeilles et al. 2011; Bobin et al. 2013, and references therein). The estimated CMB map can then be subtracted from frequency maps to reveal the weaker IR signal from unresolved sources as performed for example by Planck-Collaboration (2011).

Until now, I have considered that the CIB was the only source of non-Gaussianity. However when several non-Gaussian signals are present, a joint estimation of their amplitudes taking their covariances into account is necessary. In this section I focus on the main extragalactic non-Gaussian signals : unclustered sources (radio but more generally any population exhibiting flat moments), IR point-sources and the primordial NG of the local-type. A joint estimation then requires minimisation of the χ^2 :

$$\chi^2(f_{\text{NL}}, b_{\text{PS}}, A_{\text{IR}}) = \sum_{\ell_1 \leq \ell_2 \leq \ell_3} \frac{(\mathbf{b}_{123}^{\text{obs}} - \mathbf{b}_{123}^{\text{model}}(f_{\text{NL}}, b_{\text{PS}}, A_{\text{IR}}))^2}{\sigma^2(\ell_1, \ell_2, \ell_3)} \quad (5.19)$$

with :

$$\mathbf{b}_{123}^{\text{model}}(f_{\text{NL}}, b_{\text{PS}}, A_{\text{IR}}) = f_{\text{NL}} \mathbf{b}_{123}^{\text{CMB}} + b_{\text{PS}} \mathbf{b}_{123}^{\text{RAD}} + A_{\text{IR}} \mathbf{b}_{123}^{\text{IR}} \quad (5.20)$$

and (involved in the definition of $\sigma^2(\ell_1, \ell_2, \ell_3)$) :

$$\mathcal{C}_{\ell}^{\text{tot}} = (C_{\ell}^{\text{CMB}} + C_{\ell}^{\text{RAD}} + C_{\ell}^{\text{IR}}) B_{\ell}^2 + C_{\ell}^{\text{noise}}. \quad (5.21)$$

For backward-compatibility with the literature, I use the standard parameter b_{PS} instead of noting A_{RAD} . Note that the radio NG amplitude is heavily dependent on the flux cut through Eq.4.14. The flux cut thus needs to be specified to predict b_{PS} theoretically.

Let us define the scalar product between two bispectra \mathbf{b}^{α} and \mathbf{b}^{β} :

$$\langle \mathbf{b}^{\alpha}, \mathbf{b}^{\beta} \rangle = \sum_{\ell_1 \leq \ell_2 \leq \ell_3} \frac{\mathbf{b}_{123}^{\alpha} \mathbf{b}_{123}^{\beta}}{\sigma^2(\ell_1, \ell_2, \ell_3)} \quad (5.22)$$

Minimising Eq.5.19 corresponds to solving the linear system :

$$\begin{pmatrix} \langle \mathbf{b}_{\text{CMB}}, \mathbf{b}_{\text{CMB}} \rangle & \langle \mathbf{b}_{\text{RAD}}, \mathbf{b}_{\text{CMB}} \rangle & \langle \mathbf{b}_{\text{IR}}, \mathbf{b}_{\text{CMB}} \rangle \\ \langle \mathbf{b}_{\text{CMB}}, \mathbf{b}_{\text{RAD}} \rangle & \langle \mathbf{b}_{\text{RAD}}, \mathbf{b}_{\text{RAD}} \rangle & \langle \mathbf{b}_{\text{IR}}, \mathbf{b}_{\text{RAD}} \rangle \\ \langle \mathbf{b}_{\text{CMB}}, \mathbf{b}_{\text{IR}} \rangle & \langle \mathbf{b}_{\text{RAD}}, \mathbf{b}_{\text{IR}} \rangle & \langle \mathbf{b}_{\text{IR}}, \mathbf{b}_{\text{IR}} \rangle \end{pmatrix} \cdot \begin{pmatrix} f_{\text{NL}} \\ b_{\text{PS}} \\ A_{\text{IR}} \end{pmatrix} = \begin{pmatrix} \langle \mathbf{b}_{\text{obs}}, \mathbf{b}_{\text{CMB}} \rangle \\ \langle \mathbf{b}_{\text{obs}}, \mathbf{b}_{\text{RAD}} \rangle \\ \langle \mathbf{b}_{\text{obs}}, \mathbf{b}_{\text{IR}} \rangle \end{pmatrix} \quad (5.23)$$

If I introduce the estimators which consider only one source of non-Gaussianity, noting them with upper tilde, e.g. :

$$\tilde{A}_{\text{IR}} = \frac{\langle \mathbf{b}_{\text{obs}}, \mathbf{b}_{\text{IR}} \rangle}{\langle \mathbf{b}_{\text{IR}}, \mathbf{b}_{\text{IR}} \rangle} \quad (5.24)$$

then Eq.5.23 can be rewritten to define the joint NG estimators –noted with upper hat– as :

$$\begin{pmatrix} \hat{f}_{\text{NL}} \\ \hat{b}_{\text{PS}} \\ \hat{A}_{\text{IR}} \end{pmatrix} = \mathcal{M}^{-1} \cdot \begin{pmatrix} \tilde{f}_{\text{NL}} \\ \tilde{b}_{\text{PS}} \\ \tilde{A}_{\text{IR}} \end{pmatrix} \quad (5.25)$$

with :

$$\mathcal{M} = \begin{pmatrix} 1 & \Delta f_{\text{NL}}^{\text{RAD}} & \Delta f_{\text{NL}}^{\text{IR}} \\ \Delta b_{\text{PS}}^{\text{CMB}} & 1 & \Delta b_{\text{PS}}^{\text{IR}} \\ \Delta A_{\text{IR}}^{\text{CMB}} & \Delta A_{\text{IR}}^{\text{RAD}} & 1 \end{pmatrix} \quad (5.26)$$

and :

$$\begin{aligned} \Delta f_{\text{NL}}^{\text{RAD}} &= \frac{\langle \mathbf{b}_{\text{RAD}}, \mathbf{b}_{\text{CMB}} \rangle}{\langle \mathbf{b}_{\text{CMB}}, \mathbf{b}_{\text{CMB}} \rangle} & \Delta f_{\text{NL}}^{\text{IR}} &= \frac{\langle \mathbf{b}_{\text{IR}}, \mathbf{b}_{\text{CMB}} \rangle}{\langle \mathbf{b}_{\text{CMB}}, \mathbf{b}_{\text{CMB}} \rangle} \\ \Delta b_{\text{PS}}^{\text{CMB}} &= \frac{\langle \mathbf{b}_{\text{CMB}}, \mathbf{b}_{\text{RAD}} \rangle}{\langle \mathbf{b}_{\text{RAD}}, \mathbf{b}_{\text{RAD}} \rangle} & \Delta b_{\text{PS}}^{\text{IR}} &= \frac{\langle \mathbf{b}_{\text{IR}}, \mathbf{b}_{\text{RAD}} \rangle}{\langle \mathbf{b}_{\text{RAD}}, \mathbf{b}_{\text{RAD}} \rangle} \\ \Delta A_{\text{IR}}^{\text{CMB}} &= \frac{\langle \mathbf{b}_{\text{CMB}}, \mathbf{b}_{\text{IR}} \rangle}{\langle \mathbf{b}_{\text{IR}}, \mathbf{b}_{\text{IR}} \rangle} & \Delta A_{\text{IR}}^{\text{RAD}} &= \frac{\langle \mathbf{b}_{\text{RAD}}, \mathbf{b}_{\text{IR}} \rangle}{\langle \mathbf{b}_{\text{IR}}, \mathbf{b}_{\text{IR}} \rangle} \end{aligned} \quad (5.27)$$

Note in particular that the estimators \tilde{f}_{NL} and \tilde{b}_{PS} are the ones defined in Komatsu et al. (2005), while \tilde{A}_{IR} is the newly proposed estimator.

For illustration, assuming ERCSC flux cuts (Planck Collaboration et al. 2011b) and including the instrumental beam and noise specified Table 5.2, the mixing matrix at 220 GHz with $\ell_{\text{max}} = 2048$ is :

$$\mathcal{M} = \begin{pmatrix} 1 & 0.0461 & 0.245 \\ 1.21 \cdot 10^{-4} & 1 & 2.12 \\ 1.31 \cdot 10^{-4} & 0.435 & 1 \end{pmatrix} \quad (5.28)$$

where b_{PS} is normalised to its expected value so that all amplitude parameters are of order $\mathcal{O}(1)$. That is, the matrix is understood to be applied to $b_{\text{PS}}/b_{\text{PS}}^{\text{exp}}$, where the expected $b_{\text{PS}}^{\text{exp}}$ was computed following formula 4.14 and using number counts for radio sources by Tucci et al. (2011).

The matrix \mathcal{M} , Eq.5.28, shows, at this frequency, that the contamination by primordial NG to point-sources NG estimators is negligible (see first column). Furthermore and unless A_{IR} and $b_{\text{PS}}/b_{\text{PS}}^{\text{exp}}$ are much bigger than one, the contamination by point-sources to the primordial NG estimator is also negligible compared to its error bars ($\sigma(f_{\text{NL}}) \sim 5$). Besides, the RAD-IR submatrix, with a condition number ~ 100 , shows much more coupling between the b_{PS} and A_{IR} estimations. They are relatively degenerate. At this frequency, we see that the CIB NG is slightly dominant, by a factor ~ 2 . These conclusions however are only valid for this frequency, for local type primordial NG, and for this flux cut.

I have examined the cases at other frequencies and other flux cuts. Lower flux cuts upweights the CIB NG compared to the radio NG. That is, \mathcal{M}_{23} is increased while \mathcal{M}_{32} is decreased. Indeed IR sources are faint, so that the CIB bispectrum is mostly unaffected by the flux cut. On the contrary, the radio bispectrum is more affected, as it is dominated by sources just below the flux cut. However, this effect is less important than the frequency dependence. Indeed, the frequency behaviour of both signals, and the fact that the bispectra have a cubic dependence in the emission law, well-defined domination depending on frequency. That is, the radio NG dominates below 220 GHz while the IR dominates above, and 220 GHz is the only frequency at which their bispectra are of same order. Thus, 220 GHz is the only frequency at which \mathcal{M}_{23} and \mathcal{M}_{32} are of order $\mathcal{O}(1)$, at other frequencies this entries are either $\ll 1$ or $\gg 1$.

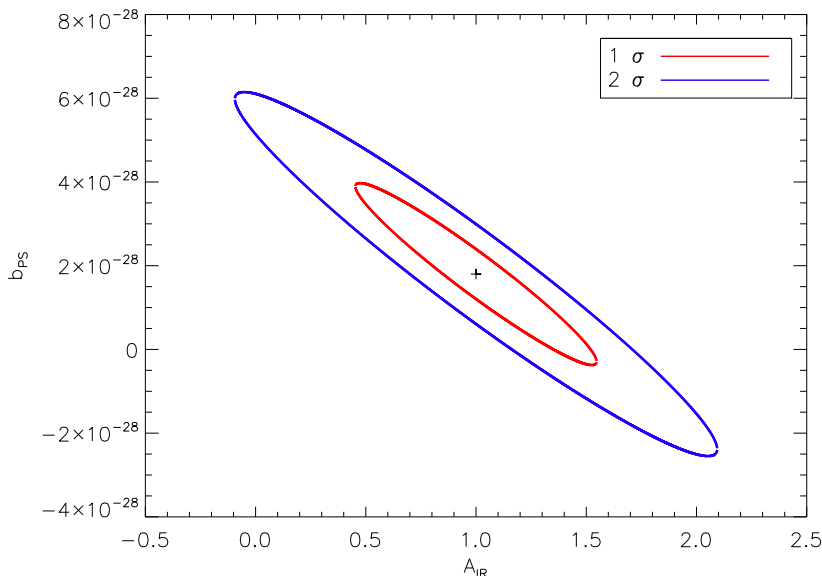


Figure 5.1: 1σ and 2σ confidence level for $(A_{\text{IR}}, b_{\text{PS}})$ in a joint estimation at 220 GHz $\ell_{\text{max}} = 2048$ with 5 arcmin Gaussian beam and $10^{-8} \Delta T/T$ noise

At 220 GHz, the primordial NG estimation decouples from the point-source estimation, however A_{IR} and b_{PS} need to be estimated jointly. For illustration, Fig. 5.1 shows the likelihood contours in the $(A_{\text{IR}}, b_{\text{PS}})$ plane, showing the expected combined constraints on these parameters at 220 GHz. A_{IR} and b_{PS} estimations are quite degenerate, as expected from the high correlations of their templates. This significantly degrades the constrain that can be put on each of them independently.

Indeed, if the radio and CIB bispectrum templates were not correlated, b_{PS} would be detected with a $\sim 3\sigma$ significance and A_{IR} would be detected with a $\sim 6.7\sigma$ significance. The combined constraint decreases these significances to a $\text{SNR} \sim 0.8$ for b_{PS} and a $\text{SNR} \sim 1.8$ for A_{IR} . There are nevertheless some possibilities to improve the constraints : including the physically-motivated prior $b_{\text{PS}} \geq 0$, and also to some extent $A_{\text{IR}} \geq 0$, and CMB removal may also help by decreasing the variance of both estimators (although not their correlation). Experimentally, the latter point is realistically achievable, and the maps I will analyse in Sect.5.2.3 & 5.3 will indeed be cleaned from CMB contamination.

5.2.2 Contamination of primordial NG estimation

As described in the previous section, the non-Gaussianity of extragalactic point-sources biases the estimation of primordial NG. This has been one of the concerns for the *Planck* NG study, whose main scientific goal is the primordial NG estimation, and is generally a concern for any NG study with CMB data. From the *Planck* point of view, the main interest was to quantify this contamination and, mostly for radio sources, to assess whether the masking of detected sources –either with the ERCSC (Planck Collaboration et al. 2011b) or with the PCCS (Planck Collaboration XXVIII 2013)– was sufficient to control it.

In this section, I will describe the different estimates of primordial NG contamina-

Radio sources						
ν (GHz)	30	90	148	219	277	350
$\ell_{\max} = 700$	108	0.17	0.0071	0.0031	0.0035	0.0064
$\ell_{\max} = 2048$	4930	7.5	0.31	0.14	0.16	0.29
CIB						
ν (GHz)	30	90	148	219	277	350
$\ell_{\max} = 700$	$-1.3 \cdot 10^{-6}$	-0.00019	-0.0033	-0.078	-0.74	-11
$\ell_{\max} = 2048$	$-1.8 \cdot 10^{-5}$	-0.0026	-0.039	-0.67	-6.3	-66

Table 5.3: From Lacasa et al. (2012), bias of the f_{NL} estimator due to radio and IR sources, measured on Sehgal et al. (2010) simulations.

tions that I have computed through different methods.

In a first step, I applied the KSW estimator to the radio and IR simulated maps by Sehgal et al. (2010). These estimates do not consider any instrumental beam nor noise. Furthermore, for these computations $\sigma^2(\ell_1, \ell_2, \ell_3)$ in Eq.5.22 only contains C_ℓ^{CMB} . That is, these computations give the bias Δf_{NL} that would arise *if the map was thought to be pure CMB*. Moreover, I performed different estimations depending on whether or not a flux cut is applied and the maximum multipole of analysis. The results are shown in Table 5.3 for radio and IR sources, in the case where the flux cut is implemented. For the flux cut, I used the ERCSC values, or the nearest value when the frequency did not match, and masked the maps at pixels above the threshold. Note that the flux cut did not have any effect for IR sources below 219 GHz, as all sources were fainter than the flux cut at these frequencies.

I also considered different maximum multipoles. The case $\ell_{\max} = 700$ is representative of *WMAP*, in this case we see that the bias is always negligible compared to the error bar $\sigma(f_{\text{NL}}) \approx 20$ (Komatsu et al. 2011) except at the frequencies 30 & 350 GHz. The case $\ell_{\max} = 2048$ is representative of *Planck*, in this case we see that the bias is often comparable to, or greater than, the error bar $\sigma(f_{\text{NL}}) \approx 5$ (Planck Collaboration XXIV 2013). In details, the bias is important for radio sources at 30 to 90 GHz, and at 277 to 350 GHz for IR sources. I have produced similar estimations without implementing the flux cut, for comparison. We then see that the flux cut greatly reduces the radio bias (I.e. the bias is much more important without flux cut), while it mostly does not affect the IR bias. This means that at high frequencies, it is important to consider the CIB NG even with a conservative masking of sources.

One caveat with these computations, are the limitations of the Sehgal et al. (2010) simulations. For example the CIB power spectrum in these simulations is too steep ($C_\ell^{\text{IR}} \sim \ell^{-2}$ while Planck-Collaboration (2011) has shown that $C_\ell^{\text{IR}} \sim \ell^{-1}$). Furthermore, the octant replication yields an artificial break of power at low multipoles, to which the KSW estimator is sensitive as local type primordial NG peaks in the squeezed configurations (where one multipole is small). Hence my estimations may be artificially biased by this simulation's shortcoming.

We now turn to analytic computation of the primordial NG bias using Eq.5.27. First, I computed the bias on f_{NL} estimation, but also its error bar, due to radio sources. This

error bar is given by :

$$\sigma^2(f_{\text{NL}}) = \left(\sum_{\ell_1 \leq \ell_2 \leq \ell_3} \frac{(b_{123}^{\text{CMB}})^2}{\sigma^2(\ell_1, \ell_2, \ell_3)} \right)^{-1} \quad (5.29)$$

analogously to Eq.5.10. The results are shown in Table 5.4, for Gaussian beams with FWHM from Planck Collaboration VII (2013), noise power spectra measured on *Planck* data, $\ell_{\text{max}} = 2048$, ERCSC flux cuts, and radio number counts by Tucci et al. (2011).

ν (GHz)	30	44	70	100	143	217	353	545	857
$\Delta f_{\text{NL}}^{\text{RAD}}$	74.5	31.9	2.93	0.82	0.28	0.096	0.074	0.63	195
$\sigma(f_{\text{NL}})$	23.5	19.4	11.4	7.70	5.63	5.48	8.62	42.2	40827

Table 5.4: $\Delta f_{\text{NL}}^{\text{RAD}}$ and $\sigma(f_{\text{NL}})$ at *Planck* frequencies, considering only unresolved radio sources. Using Gaussian beams with FWHM from Planck Collaboration VII (2013), noise power spectra measured on *Planck* data, $\ell_{\text{max}} = 2048$, ERCSC flux cuts, and radio number counts by Tucci et al. (2011).

We see that the bias is under control except at the extreme *Planck* frequencies 30-44-857 GHz, in particular the bias is negligible at the central CMB frequencies 100-143-217 GHz, where the error bars on f_{NL} are the smallest. Note that the numbers in Table 5.4 would change with a more aggressive masking, e.g. with PCCS flux cuts (Planck Collaboration XXVIII 2013), although the main conclusions would be unchanged.

Finally, I have computed mixing matrices, described in Sect.5.2.1, for *Planck* HFI frequencies. For the computation of these matrices, I used the same instrumental and radio characteristics as for Table 5.4, and for the CIB I used power spectra (needed for the prescription for b_{123}^{CIB} and for the error bars $\sigma^2(\ell_1, \ell_2, \ell_3)$) provided by the CIB model described in Planck-Collaboration (2011) and reproducing the *Planck* measurement (*ibid.*). I do not present all matrix elements here but just those associated with the bias on f_{NL} estimation. They are shown in Table 5.5, along with the correlation coefficient $r(\text{IR}, \text{RAD})$, between the radio and IR bispectra, defined as :

$$r(\text{IR}, \text{RAD}) = \frac{\langle \mathbf{b}_{\text{IR}}, \mathbf{b}_{\text{RAD}} \rangle}{\sqrt{\langle \mathbf{b}_{\text{IR}}, \mathbf{b}_{\text{IR}} \rangle \langle \mathbf{b}_{\text{RAD}}, \mathbf{b}_{\text{RAD}} \rangle}} \quad (5.30)$$

Note that Table 5.5 only considers *Planck* HFI frequencies, as the CIB signal is completely negligible at LFI frequencies.

One can in particular compare the obtained $\Delta f_{\text{NL}}^{\text{RAD}}$ with that shown in Table 5.4. We see that they are in agreement at 100 & 143 GHz and start gradually disagreeing afterwards, until having opposite sign at 857 GHz. The reason for this is that in Table 5.4 only radio sources were considered, so that $\sigma^2(\ell_1, \ell_2, \ell_3)$ contained C_{ℓ}^{CMB} and C_{ℓ}^{RAD} , while for Table 5.5 we also have a contribution of C_{ℓ}^{IR} to $\sigma^2(\ell_1, \ell_2, \ell_3)$. Therefore the estimates are in agreement at low frequencies, where C_{ℓ}^{IR} is negligible, and start to disagree as frequency increases.

The sign reversals of $\Delta f_{\text{NL}}^{\text{RAD}}$ at 857 GHz between Table 5.4 and 5.5, and of $\Delta f_{\text{NL}}^{\text{IR}}$ between 353 and 545 GHz show the sensitivity of the biases to the conditions at high multipoles.

ν (GHz)	100	143	217	353	545	857
$\Delta f_{\text{NL}}^{\text{RAD}}$	0.82	0.27	0.089	0.046	0.016	-391
$\Delta f_{\text{NL}}^{\text{IR}}$	$8.2 \cdot 10^{-4}$	$5.6 \cdot 10^{-3}$	0.11	19.3	-1749	$-7.43 \cdot 10^7$
$r(\text{IR}, \text{RAD})$	94.3%	94.8%	94.8%	95.2%	92.8%	91.9%

Table 5.5: Bias of the f_{NL} estimator for local type primordial NG due to radio and IR sources, and correlation between the radio and IR bispectrum templates, computed with the mixing matrices described in Sect.5.2.1. Using Gaussian beams with FWHM from Planck Collaboration VII (2013), noise power spectra measured on *Planck* data, $\ell_{\text{max}} = 2048$, ERCSC flux cuts, and radio number counts by Tucci et al. (2011).

Indeed, the radio and IR power spectrum become dominant over the CMB at high multipoles, hence the f_{NL} estimator downweights these configurations. When this domination extends to lower multipoles, e.g. when C_{ℓ}^{IR} is added to C_{ℓ}^{RAD} or at high frequencies, the configurations with low multipoles are favoured, and the CMB is negative in these configurations.

The same reasoning also shows the importance of the noise model. Indeed, the noise power spectrum used for Table 5.4 and 5.5 is not flat but decreasing with multipoles. In particular, its value in the first 100 multipoles is much larger than at higher multipoles where it flattens. This changes the weighting of bispectrum configurations in the estimators and in the matrix elements Eq.5.27, particularly in the squeezed limit. Indeed, assuming a white-noise changes the values of $\Delta f_{\text{NL}}^{\text{RAD}}$ and $\sigma(f_{\text{NL}})$, and may induce for example a sign reversal of $\Delta f_{\text{NL}}^{\text{RAD}}$ at 857 GHz.

Because of the oscillations and sign-changes of b_{123}^{CMB} , the radio and CMB bispectra are nearly orthogonal, and the IR and CMB bispectra also. Hence, some changes of the weighting of bispectrum configurations in the estimators may easily change the value and sign of the scalar product. This also explains how, at 545 GHz in Table 5.5, $\Delta f_{\text{NL}}^{\text{RAD}}$ can be positive while $\Delta f_{\text{NL}}^{\text{IR}}$ is negative, even if the bispectra are $\sim 93\%$ correlated, as illustrated schematically in Fig.5.2

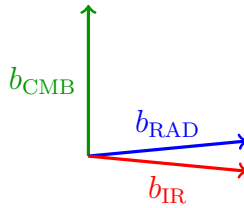


Figure 5.2: Schematic illustration of how $\Delta f_{\text{NL}}^{\text{RAD}}$ can be positive while $\Delta f_{\text{NL}}^{\text{IR}}$ is negative. Indeed, the bispectrum vectors b_{IR} and b_{RAD} are nearly colinear, but the scalar product $\langle b_{\text{RAD}}, b_{\text{CMB}} \rangle$ is positive while $\langle b_{\text{IR}}, b_{\text{CMB}} \rangle$ is negative.

The *Planck* NG article (Planck Collaboration XXIV 2013) has used the *Planck* data to put constraint on the different sources of non-Gaussianity in the *Planck* CMB maps. To this end, different CMB maps produced by several component separation methods have been considered, as well as the raw 143 GHz channel. Several estimators based on the 3-point function (except for Minkowsky functionals which test all orders) have been applied to the data and extensively tested on realistic Gaussian and non-Gaussian

simulations. Furthermore, a joint NG estimation has been used considering the following sources of non-Gaussianity : primordial NG of the local, equilateral and orthogonal types (see Sect.3.3); b_{PS} template; and ISW-lensing (see Sect.3.3 and 4.1).

It can be noted that the CIB template was not considered. This choice was made because of the following considerations : time constraints, the high correlation ($> 90\%$) between the radio and CIB templates, and the fact that the maps used were expected by the working group to be more contaminated by radio than by IR sources (in particular the 143 GHz channel). Furthermore, the robustness of PNG estimations to point-sources is also ensured by the fact that the obtained f_{NL} values do not change much if b_{PS} is not considered in the joint estimation (at most by 0.3σ).

However, the obtained b_{PS} estimation can difficultly be used for point-source constraints for the following reasons. First, it is potentially a mix of b_{PS} and A_{IR} as already outlined. Then it was obtained with a conservative masking (union of sources detected at any frequency) which does not translate into a flux cut or a single selection function. Finally, except for the 143 GHz value, it was obtained on component separated maps which are designed to decrease the foregrounds signals and whose exact effect on a point-source bispectrum is unknown.

For the Cosmic Infrared Background, a dedicated NG study is thus necessary to use the information present in the *Planck* maps, particularly above 217 GHz.

5.2.3 Measuring the CIB NG with *Planck* data

This section describes the *Planck* data analysis performed to detect the CIB bispectrum, as part of the results published in Planck Collaboration XXX. (2013).

Let us first describe the *Planck* CIB article : the study is based on the *Planck* maps as well as the IRAS maps at 3000 GHz (Neugebauer et al. 1984) reprocessed by Miville-Deschênes & Lagache (2005). For the *Planck* channels, two maps are used per frequency, constructed respectively from the first and last rings (see Sect.5.1). The goal of the article is to measure the CIB anisotropies and model them so as to constrain the star formation history of the universe. To this end, the CIB anisotropies are unveiled through a cleaning of the maps detailed below. The power spectrum analysis proceeds to estimate the maps spectra and cross-spectra through cross-correlation of the first and last maps. Some contaminations are further accounted for at the power spectrum level (tSZ and CIB leakage, see sect.5.2.3.3), before a halo model is fitted to the data and the results are interpreted in term of galaxy bias and star formation history.

To extract the CIB from the *Planck* maps, two components need to be removed : the CMB and the Galactic dust. The CMB was removed by subtracting an estimated CMB map obtained through Wiener filtering of the *Planck* 100 GHz map :

$$a_{\ell m}^{\text{corr}}(\nu) = a_{\ell m}(\nu) - w_{\ell} a_{\ell m}(100 \text{ GHz}) \quad (5.31)$$

with w_{ℓ} the 100 GHz Wiener filter.

This simple technique allows propagation of the errors and of contaminations (by the CIB and tSZ present in the 100 GHz map in particular). On the contrary, for a CMB map estimated by multi-frequency component separation, the contamination is unknown¹.

The Galactic dust component was removed by constructing a dust model with atomic

¹In particular the contamination of the estimated CMB by the CIB itself is problematic.

hydrogen (HI) observations, as Boulanger et al. (1996) reported a tight dust-HI correlation except for high HI column densities. Two different sets of HI observations were used for the studies I will present. First, HI observations by the Green Bank Telescope (Boothroyd et al. 2011) have a ~ 9.5 arcmin resolution and define nine fields with a total area of $\sim 0.6\%$ of the sky. Second, Parkes telescope observations define the GASS HI survey (McClure-Griffiths et al. 2009), with much larger sky coverage ($\sim 10\%$ of the sky) albeit with a lower (16.2 arcmin) resolution.

5.2.3.1 Small fields

The first fields I used for the bispectrum analysis were small patches (hence hereafter called small fields) cleaned with the high-resolution Green Bank Telescope HI observations. There are nine fields, with historical names : N1, AG, SP, NEP4, SPC4, SPC5, LH2, MC and Bootes. Their area ranges from 0.04 to 0.1% of the sky, with a total area $\sim 0.6\%$ of the sky. For illustration, Fig.5.3 shows three of the fields at the frequencies used for the power spectrum analysis (i.e., *Planck* 143-217-353-545-857 GHz and IRAS 3000 GHz).

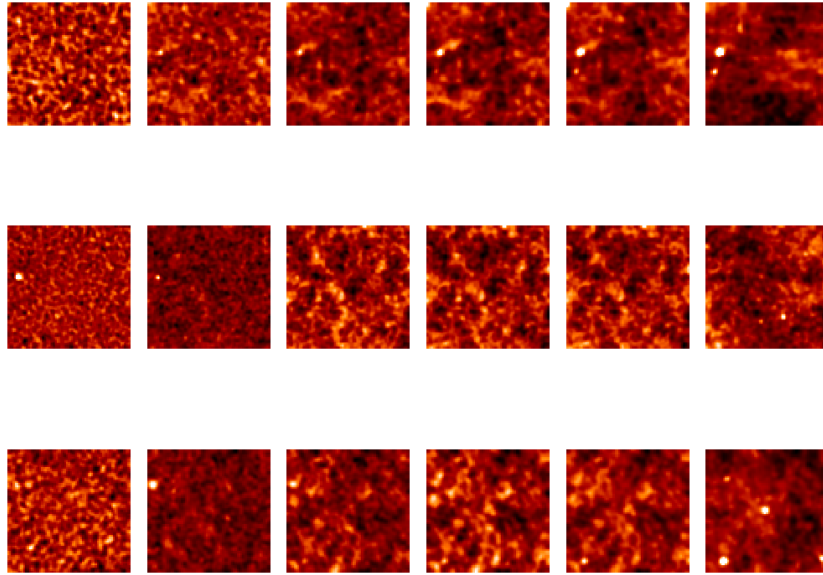


Figure 5.3: Three of the cleaned small CIB fields. *From top to bottom* : SPC4, N1 and LH2. The maps at the different frequencies are shown, *from left to right* : *Planck* 143-217-353-545-857 GHz and IRAS 3000 GHz.

To use these patches, I developed a flat sky bispectrum estimator. Fig.5.4 shows the resulting bispectra obtained for each field, at 545 GHz.

We see a large dispersion between the bispectrum estimates, as they are variance-dominated due to the small sky coverage. The estimates however tend to be positive and larger at low multipoles than at high multipoles. To obtain a single estimate with a lower variance, I have computed an inverse-variance weighted average of the bispectra.

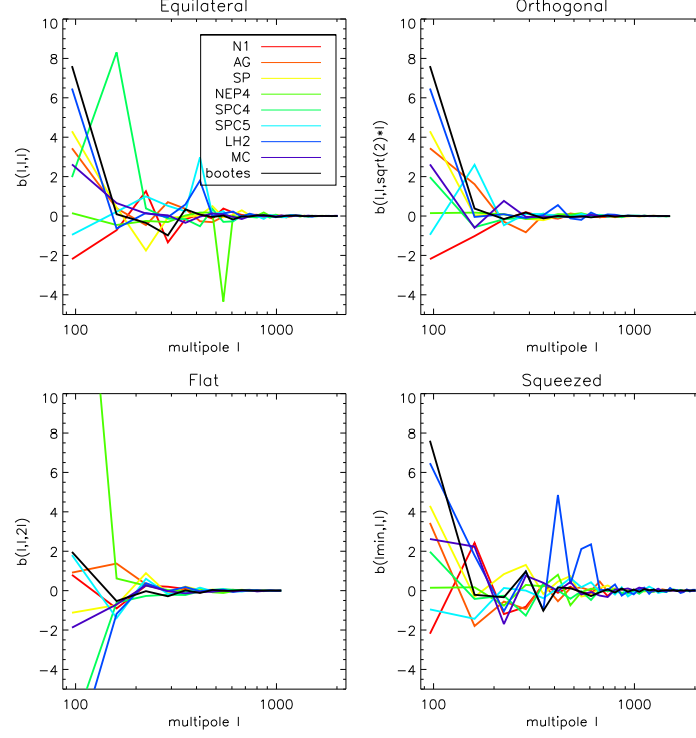


Figure 5.4: Bispectra of the nine different small CIB fields at 545 GHz.

More precisely, we have seen in Sect.2.2.5.2 that, in the weak NG case, the bispectrum variance is :

$$\text{Var}(b_{123}) = \frac{C_{\ell_1} C_{\ell_2} C_{\ell_3}}{N_{123}} \Delta_{\ell_1 \ell_2 \ell_3} \quad (5.32)$$

In the flat sky case, the expression is analog but N_{123} is replaced by $N(k_1, k_2, k_3)$ the number of triplets of Fourier modes such that $\vec{k}_1 + \vec{k}_2 + \vec{k}_3 = \vec{0}$. This latter number depends on the size and geometry of the field considered : the smaller the field, the fewer Fourier modes. Hence the inverse-variance weighting of the bispectra corresponds to a $N(k_1, k_2, k_3)$ weighting. The resulting average bispectrum at 217 GHz is shown in Fig.5.5 along with the expected CIB bispectrum (computed with the prescription) for comparison.

We see that the inverse-variance weighted average has importantly reduced the variance of the bispectrum estimate which seems to follow the prescription at the order of magnitude level. However, the estimate is still too noisy because of the limited sky coverage. In the next section, I describe the analysis of a larger field.

5.2.3.2 GASS field : bispectrum estimation pipeline

The other field I have analysed was cleaned with the GASS HI survey (McClure-Griffiths et al. 2009). An estimated map of the Galactic dust emission was constructed with this survey, as detailed in Planck collaboration & GASS collaborators (2013); Planck Collaboration XXX. (2013). This estimated map will be referred to as the dust model hereafter. Compared to the small fields described in the previous section, the GASS field

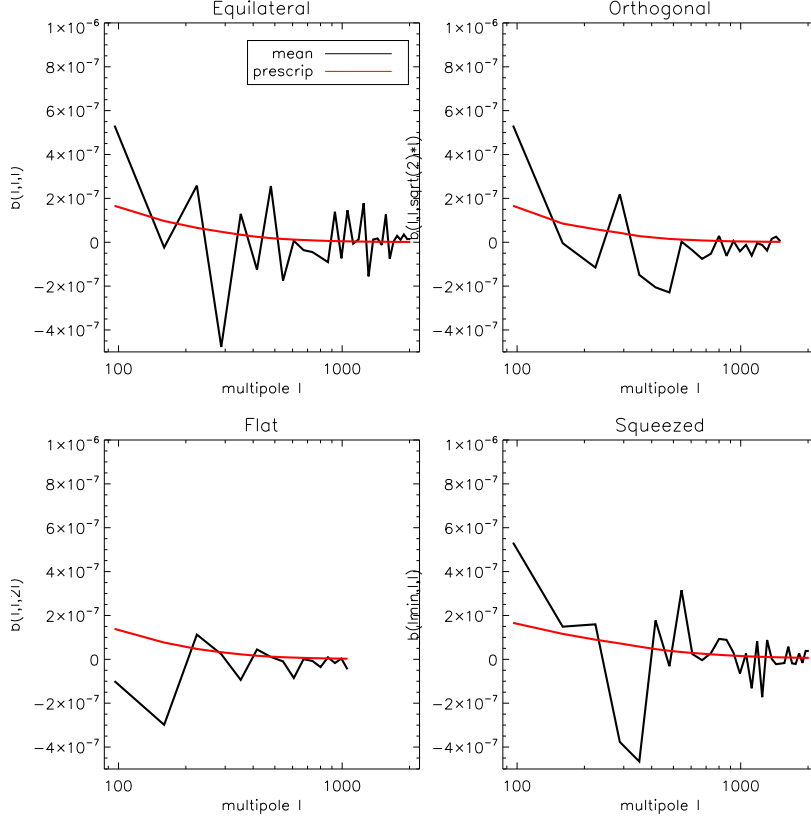


Figure 5.5: $1/f_{\text{SKY}}$ weighted mean of the small CIB fields bispectra at 217 GHz (black), with the empirical prescription for the bispectrum overplotted in red for comparison.

provides a much larger sky fraction, $\sim 10\%$ of the sky, but with lower resolution (16.2 arcmin). For illustration, figure 5.6 shows the cleaned CIB map obtained by Planck Collaboration XXX. (2013) at 217 GHz.

The map is centered around the Galactic South pole and at $N_{\text{side}} = 512$ HEALPix resolution, so that the maximum multipole of analysis is $\ell_{\text{max}} = 1024$. With this resolution, the computation of a bispectrum is relatively fast which allows us to easily run tests and simulations.

The GASS field is large, but a complex mask is needed as seen in Fig.5.6. As shown in Sect.2.2.4.4 and 2.2.4.5, this may induce a bias to the bispectrum estimation, and the coupling matrix cannot be computed at this resolution. To tackle this issue, I first chose a binning size large enough to reduce the coupling between bins, but not too large so as not to lose information on the variations of the bispectrum. I computed the power spectrum coupling matrix to estimate the necessary minimal bin size and found two choices : $\Delta\ell = 64$ and $\Delta\ell = 128$. Since the estimates with $\Delta\ell = 64$ were still noisy (SNR per configuration lower than one), I chose $\Delta\ell = 128$. Indeed I wanted to be able to visualise the bispectra, to detect possible contaminations and problematic configurations. I checked that this bin size did not significantly lose bispectrum features². Then I built

²With the power-law model found in Sect.5.2.3.4, I computed that there is a $> 96\%$ correlation between the binned and the full bispectra

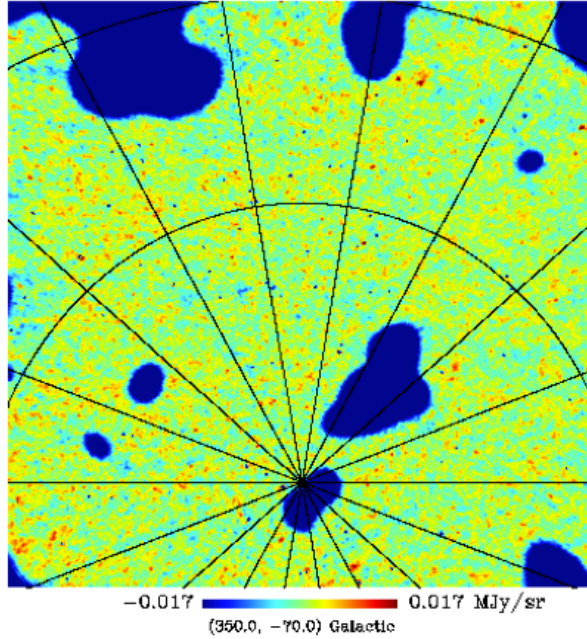


Figure 5.6: From Planck Collaboration XXX. (2013), cleaned CIB map at 217 GHz in the GASS field.

an average debiasing of the bispectra, assuming a template for it. To do so, I have generated maps with a bispectrum given by the empirical prescription, smoothed them at the GASS resolution (16.2 arcmin), masked them, measured the masked bispectrum and compute the ratio masked/full-sky in each configuration, and average this ratio over simulations. Figure 5.7 shows this ratio averaged over 50 simulations, which were enough to reach convergence. It also shows the small error bars on the estimation of this ratio and the analytical ratio in the case of the f_{SKY} approximation.

We see that the measured ratio follows very closely the f_{SKY} approximation indicating that the bin-coupling is small and that we have no complex mask effect. However, it is worth emphasising that this result relies on the assumption that the underlying bispectrum is properly described by the empirical prescription. In the following, I will debias CIB bispectrum estimates with this ratio, and this will provide unbiased estimates as long as the empirical prescription describes correctly the CIB NG.

As shown in Sect.2.2.4.3, a linear correction to the bispectrum is normally also needed to reduce the variance of the bispectrum estimate. I have computed this linear term with 50 simulations and found it completely negligible compared to the cubic term. Indeed, the linear term is meant to correct for the coupling between different bins, but the large bin size decreases this bin coupling.

Another problem to address in the CIB analysis is the presence of dust residuals. Indeed, the dust model used to clean the *Planck* maps is not perfect, and there are residuals of Galactic dust in the maps, due in particular to “dark gas” not traced by HI (Planck Collaboration et al. 2011c). Galactic dust has a highly non-Gaussian, and non-isotropic, distribution (Planck Collaboration et al. 2011a). For example, I have computed the bispectrum of the dust model map introduced previously, it is shown in Fig.5.8 at 857 GHz. The figure also shows the empirical prescription for the CIB bispectrum for

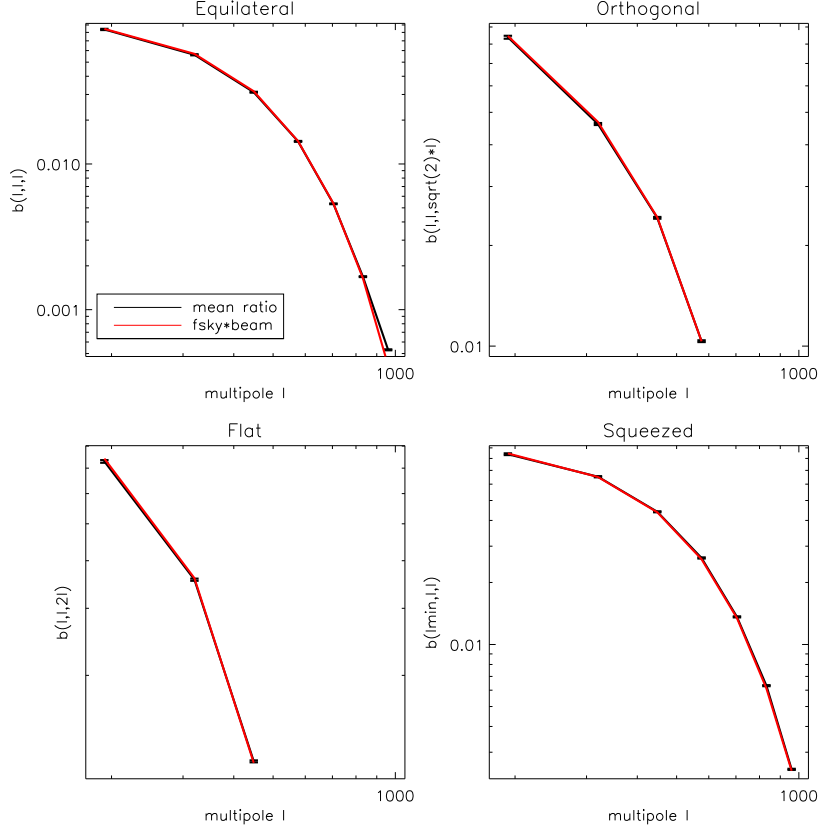


Figure 5.7: Ratio of the masked bispectrum to the full-sky bispectrum (black), compared with the f_{SKY} approximation (red).

comparison.

We see that the dust bispectrum dominates the expected CIB bispectrum on large angular scales. In particular in this plot, the first bin ($\ell = 2 \cdots 128$) where the dust bispectrum is too important was discarded. In all my CIB bispectrum estimations, due to the high dust signal as compared to the CIB, this first bin has been discarded. In the second bin, we see that the dust bispectrum must be reduced by a factor ~ 100 to uncover the CIB signal, which means that the dust model must account for $\sim 80\%$ of the dust. Planck Collaboration XXX. (2013) did estimate that there are $\sim 20\%$ dust residuals in the maps compared to the dust model, by comparing the obtained power spectrum with different dust masks. I obtained the same estimate by comparing the raw map bispectrum (without dust cleaning) and the dust model bispectrum at 857 GHz. Hence, the bispectrum estimates involving the second bin may be contaminated by dust, and I include this effect in the error bars of the CIB bispectrum estimation. We see however that CIB bispectrum estimates are safe from dust contamination at higher multipoles, as the dust bispectrum has a steep decrease.

A possibility to minimise the contamination by dust residuals was to use a mask leaving a smaller sky area ($f_{\text{SKY}}=4.6\%$) that was designed for the power spectrum analysis. After extensive testing I found that the results were not better than with $f_{\text{SKY}}=10\%$ and did not justify losing sky area and increasing cosmic variance.

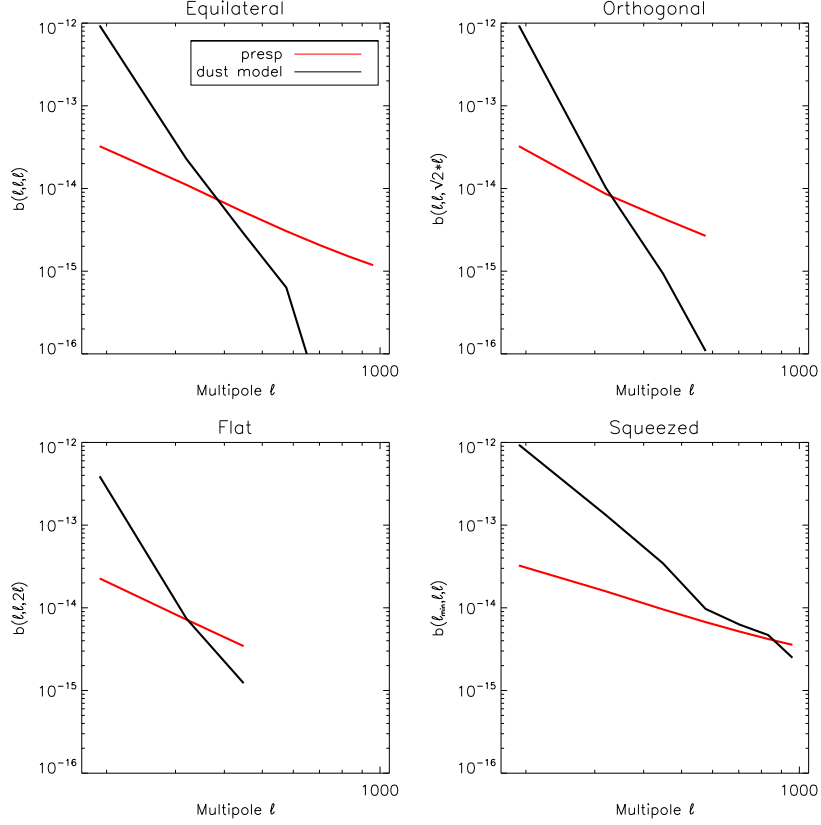


Figure 5.8: Bispectrum of the Galactic dust model at 857 GHz, compared with the empirical prescription for the CIB bispectrum.

As already mentioned, there are two *Planck* maps per frequency for the CIB analysis, respectively built from the first and last rings. The CIB power spectrum was estimated with the cross power spectrum of those maps, as this debiases the measurement from noise contamination (as the noise in the first rings is not correlated with that in the last rings, note however that the noise still increases the variance of the measurement). Since the noise is Gaussian, we do not need such considerations for the bispectrum. So I have computed auto-bispectra for the first and last rings respectively, and this allowed me to perform consistency checks by looking at their difference.

5.2.3.3 GASS field : results and post-treatment

Considering all the aforementioned details, Fig.5.9 shows the estimates of the CIB bispectrum at 545 GHz, along with the empirical prescription for comparison.

We see that the first and last-ring estimates are consistent, except possibly at high multipoles in the equilateral configuration. Moreover, they have the same order of magnitude as the prescription and a similar slope, but they are nevertheless lower. I have then combined these estimates and computed error bars, in fact a covariance, for the measurement. As described in Planck Collaboration XXX. (2013), the covariance contains the following terms :

- cosmic variance : $C_{2 \times 2 \times 2}$ term with the f_{SKY} approximation. For the power spectra

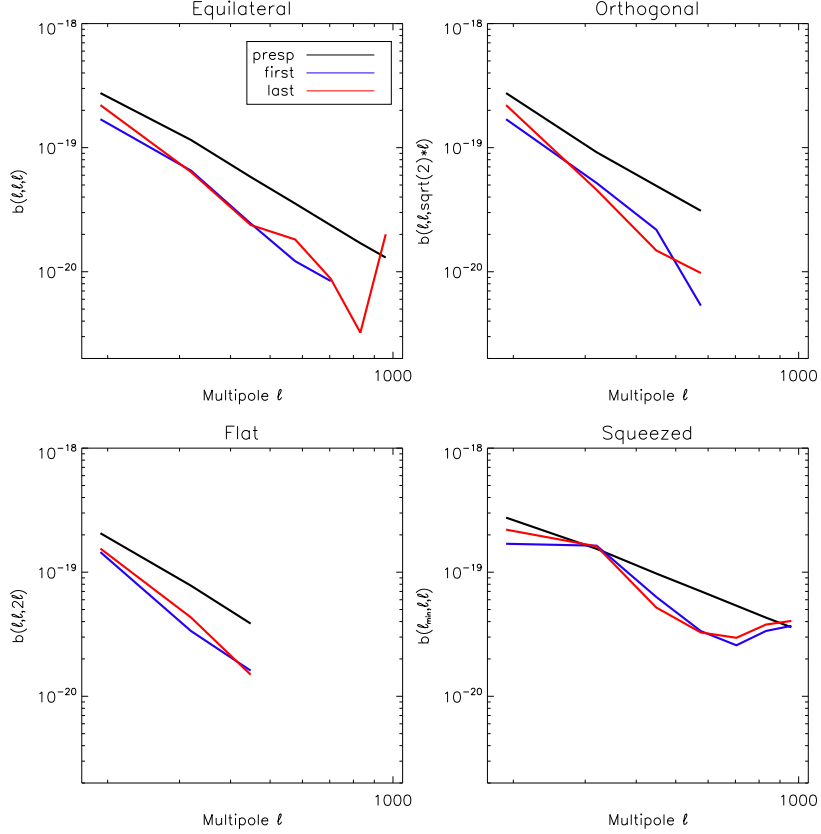


Figure 5.9: Estimated CIB bispectrum at 545 GHz with the first and last rings, along with the prescription for comparison.

involved, I estimated the first and last-ring auto-spectra, containing the noise power spectrum, as necessary.

- cosmic variance : $C_{3 \times 3}$ term with the f_{SKY} approximation. The bispectrum involved is the estimated CIB bispectrum.
- $C_{4 \times 2}$ and C_6 terms of the cosmic variance : I have neglected these terms since I did not have estimates of the CIB trispectrum nor 6-point function.
- dust residual : I used the estimated dust residual fraction and the bispectrum of the dust model. Conservatively, I estimated the bispectrum on the whole GASS field and with the two masks previously described, and took the maximum estimate per configuration.

The 143 GHz measurement is dominated by noise : the bispectrum has positive and negative values, with absolute value up to ten times bigger than the prescription prediction. The prescription values are within the error bars, but the noise is too large for the measurement to be of any interest. This is not surprising as the 143 GHz analysis was already difficult at the power spectrum level (Planck Collaboration XXX. 2013). Therefore, I did not consider this frequency any further.

I also discarded the 857 GHz frequency for CIB analysis. In this case the bispectrum

coefficients involving the second bin are negative. In the following bins, the bispectrum follows the prescription, but exceeds it and rises at high multipoles. I attribute this behaviour to dust contamination. It is indeed found by computing the bispectrum of the masked dust model and of simulations with an input bispectrum as steep as the dust one. Although the f_{SKY} approximation works well for a “shallow” bispectrum as shown in Fig.5.7, for a steep bispectrum the mask can yield a leakage of large-scale power into the small-scales. That is, even if the bin coupling is small, the large-scale power is so important that it contaminates significantly smaller scales. The measurements in the 857 GHz channel have nevertheless allowed me to pinpoint the configurations which are the most prone to dust contamination, and to have an estimation of the contamination pattern.

The CIB bispectrum measurement at 353 GHz with its error bars is shown in Fig.5.10, along with some fits described hereafter.

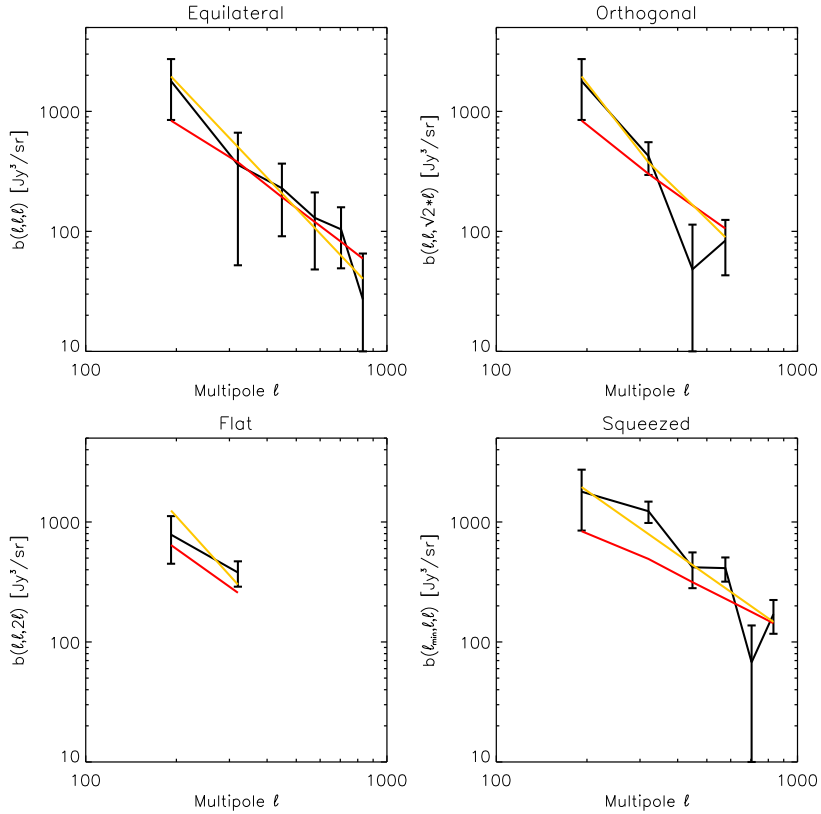


Figure 5.10: From Planck Collaboration XXX. (2013), measured CIB bispectrum at 353 GHz, along with the best-fit prescription (red) and power law (orange).

The measurements at 217 and 545 GHz show a similar behaviour, although with larger error bars at 217 GHz and smaller ones at 545 GHz.

I then proceeded to check whether the dust contamination pattern visible at 857 GHz was present in the measurements at 217 353 and 545 GHz, and to discard the configurations that seemed affected, depending on the frequencies. I found that some configurations were indeed affected, mostly at high multipoles. There were a bit more of them at 545 GHz than at 217 and 353 GHz. This result is expected as the dust emission falls more

steeply than the CIB one when frequency decreases, so that the dust contamination is more important at high frequencies. I also discarded configurations where the bispectrum was lower than a minimum acceptable value³, and where the ratio shown in Fig.5.7 was too different from the f_{SKY} approximation. However, this latter condition did not result in any exclusion, as the f_{SKY} approximation performs well in all configurations of this analysis.

Nevertheless, the measured bispectra need an additional correction. Even if they are free from CMB and Galactic dust contamination, the bispectra measure the non-Gaussianity of all the other foregrounds. Namely, there is contamination from radio and tSZ, at the observed frequency but also leaking from 100 GHz. Indeed, as stated in the introduction of this section, the CMB component is cleaned with a Wiener-filtering of the 100 GHz map. This 100 GHz map contains contributions from the radio tSZ and CIB signals which will leak into the frequency CIB maps. Hence the observed spherical harmonic coefficients take the form :

$$a_{\ell m}(\nu) = a_{\ell m}^{\text{CIB}}(\nu) + a_{\ell m}^{\text{RAD}}(\nu) + a_{\ell m}^{\text{tSZ}}(\nu) - w_{\ell} [a_{\ell m}^{\text{CIB}}(100) + a_{\ell m}^{\text{RAD}}(100) + a_{\ell m}^{\text{tSZ}}(100)] \quad (5.33)$$

where w_{ℓ} is the Wiener filter.

First, the radio contamination is negligible. The b_{PS} value for the CIB fields computed with the Tucci et al. (2011) number counts is much lower than the CIB measurements. Indeed, as the GASS field is at high galactic latitudes, the point-sources are detected down to lower fluxes than on the rest of the sphere and are thus more efficiently masked. Then, I neglected the leakage from CIB at 100 GHz. Indeed, the IR emission decreases quickly with wavelength so that the intensity at 100 GHz is much smaller than at 217 GHz, and even more so at higher frequencies. Including the CIB leakage would be a small correction, as for the power spectrum (Planck Collaboration XXX. 2013), which dominant term would be the $100 \times 217 \times 217$ cross-bispectrum.

Finally, the tSZ leakage leads to a contamination by the tSZ bispectrum and by the tSZ-CIB cross bispectra. At the power spectrum level, the cross tSZ-CIB correction is found negligible compared to the tSZ power spectrum correction (3% vs 15%, see Planck Collaboration XXX. 2013). For the bispectrum analysis, I neglected the cross-bispectra and focused on the correction of the tSZ bispectrum. For this purpose, I used the bispectrum measured on the projected tSZ catalogue (see Sect.5.3 and Planck Collaboration XXIX 2013) as a template so that :

$$b_{123}^{\text{corr-tSZ}}(\nu) = (g_{\nu} - g_{100} w_{\ell_1})(g_{\nu} - g_{100} w_{\ell_2})(g_{\nu} - g_{100} w_{\ell_3}) b_{123}^{\text{template}} \quad (5.34)$$

where g_{ν} is the factor converting Compton parameter y to μK_{CMB} at that frequency. The tSZ correction is negligible at 353 and 545 GHz, but important at 217 GHz. It is worth noting that the tSZ signal does not exactly vanish at 217 GHz. g_{217} is non-zero when we account for the *Planck* bandpass⁴. Accounting for this effect produces a $\sim 15\%$ correction to $b_{123}^{\text{corr-tSZ}}$.

The last element of the measurement is the assessment of the error bar on the tSZ subtraction. For this, I considered a 80% error bar on the overall amplitude of the bispectrum template, motivated by the difference between the bispectrum produced by

³The minimum acceptable value was 1% of the minimum value of the empirical prescription bispectrum.

⁴The signal integrated over the frequency response does not vanish exactly.

the tSZ catalogue, the measured tSZ bispectrum and that of the *Planck* tSZ simulation (see Sect.5.3 for details). The error bars corresponding to this tSZ correction are totally correlated (the corresponding covariance matrix has rank 1). I added this covariance matrix to the previously mentioned one. The total covariance matrix thus becomes highly non-diagonal at 217 GHz.

Finally, we can now assess the significance of the measurement at the three frequencies 217, 353 and 545 GHz. To this end, I have computed the mean SNR per configuration and the total SNR, which show the significance of the detection. The results are reported in Table 5.6. Note that $\text{SNR}_{\text{tot}}^2 \neq n_{\text{config}} \times \text{SNR}_{\text{config}}^2$, as the covariance matrix is non-diagonal, and as it is the mean $\text{SNR}_{\text{config}}$ not the rms which is reported. Note also that there were originally 43 configurations before discarding/flagging the contaminated ones.

frequency	number of config.	mean SNR per config.	total SNR
217 GHz	38	1.24	5.83
353 GHz	40	2.85	19.27
545 GHz	36	4.59	28.72

Table 5.6: From Planck Collaboration XXX. (2013), detection significance of the CIB bispectra at each frequency.

We find a very significant detection of the CIB bispectrum, even per configuration. Note that the significance increases with frequency as the noise becomes lower relative to the signal. These measurements can be compared with a recently announced detection of the CIB non-Gaussianity by Crawford et al. (2013) with South Pole Telescope (SPT) data. They computed a template fitting of the bispectrum measured on SPT data at 95-150-220 GHz, with three components : a b_{PS} template, a template for the CIB derived from the empirical prescription of Lacasa et al. (2012) and a template for the tSZ bispectrum derived from Bhattacharya et al. (2012). They reported a $\sim 5\sigma$ detection of the CIB bispectrum with the multi-band analysis.

In comparison with the SPT analysis, the measurements with *Planck* data provide a detection at each frequency instead of a multi-band analysis. The measurements have a higher significance, the least-significant being a 5.8σ detection, while the most significant is a 28.7σ detection. Also, I considered the non-Gaussian $C_{3 \times 3}$ contribution to the covariance, while the SPT analysis only considers the Gaussian contribution, i.e., a variance but no off-diagonal terms. As will be argued in Sect.5.3, this probably overestimates their detection significance, in particular because of the tSZ contribution in their multi-band analysis, as the tSZ is highly non-Gaussian.

5.2.3.4 GASS field : exploitation of the measurements

I compared in Fig.5.9 the bispectrum measured on the CIB maps with the empirical prescription, described in Sect.4.3.2. With the amplitude parameter $\alpha \approx 310^{-3}$, the value found on Sehgal et al. (2010) simulations, the prescription overestimates the measurement at all frequencies. So I fitted for α through a χ^2 minimisation with the covariance matrix described previously. The results are found in Table 5.7, along with the error bar on α , the best-fit χ^2 and the number of degrees of freedom.

frequency	α	χ^2	d.o.f.
217 GHz	$(1.89 \pm 0.53) \cdot 10^{-3}$	21.4	37
353 GHz	$(1.21 \pm 0.07) \cdot 10^{-3}$	45.5	39
545 GHz	$(1.56 \pm 0.06) \cdot 10^{-3}$	95.9	35

Table 5.7: From Planck Collaboration XXX. (2013), best-fit amplitude parameter for the empirical prescription, its error bar and the chi-square value of the fit.

Note that the obtained values of α are consistent with each other. This is a strong indication that the CIB measurements have little to no residual contamination. Indeed, the intrinsic level of non-Gaussianity is found consistent between frequencies, as expected. Note also that the χ^2 degrades as frequency increases because the noise decreases with frequency. At 217 GHz, the error bars may be overestimated, in particular due to the tSZ correction, so that the χ^2 is underestimated.

The prescription fit does not provide exceptional χ^2 values, furthermore the measurement seem to have a constant slope in each configuration and to be steeper than the prescription, see Fig.5.9 and 5.10. I have thus also fitted for a separable power law :

$$b_{123}^{\text{pl}} = A \times \left(\frac{\ell_1 \ell_2 \ell_3}{\ell_0^3} \right)^{-n} \quad (5.35)$$

where the amplitude A and the index n are the two parameters. I chose a pivot scale $\ell_0 = 320$, which is the center of my third multipole bin⁵. I have fitted for A and n through χ^2 minimisation with the aforementioned covariance matrix, and I computed their error bars and their correlation with a Fisher matrix. I checked that the Fisher matrix reproduces adequately the 1-2-3 σ contours of the likelihood. The results are shown in Table 5.8 : best-fit amplitude and index, their errors and correlation, and the best-fit χ^2 and numbers of d.o.f..

frequency	A [Jy ³ /sr]	n	correlation	χ^2	d.o.f.
217 GHz	$(1.46 \pm 0.68) \cdot 10^1$	0.822 ± 0.145	81.4%	20.6	36
353 GHz	$(5.06 \pm 0.49) \cdot 10^2$	0.882 ± 0.070	82.3%	34.3	38
545 GHz	$(1.26 \pm 0.09) \cdot 10^4$	0.814 ± 0.050	85.4%	82.8	34

Table 5.8: From Planck Collaboration XXX. (2013), best fit amplitude A and index n of a power law fit to the bispectra, their error bars and associated correlation, and the chi-square value of the fit. The pivot scale is $\ell_0 = 320$

We see that the indices are consistent with each other : $n \sim 0.8 - 0.9$, indicating again that there is little or no residual contamination at these frequencies. Note that the power law provides a significantly better fit to the data at 353 and 545 GHz than the prescription with lower χ^2 values. Interestingly, we find indication that the CIB bispectrum is steeper than the prescription. The obtained indices are significantly larger than those of the prescription (which predicts $n \sim 0.6$, as $C_\ell \propto \ell^{-1.2}$), which is an interesting

⁵Thus, A is the best-fit value for the CIB bispectrum in the equilateral configuration (320,320,320). A change of pivot scale amounts to a rescaling of A .

result.

Indeed, we have seen that the modelisation with the halo model, Sect.4.3.3, produces a bispectrum in general agreement with the prescription for the emissivity models considered. A possible solution to the disagreement between the halo model and the measurement would be to change the IR emissivities. For example, Béthermin et al. (2013) recently described a model for the CIB power spectrum different from what is presented in Sect.4.3.3, considering IR emissivities depending on the halo mass (and on the galaxy type : central/satellite). They showed that the IR emissivity decreases in massive haloes due to the quenching of star formation. This leads to a decrease of the power spectrum 1-halo term, so that the 2-halo term dominates over the multipole range of interest for *Planck*. This was confirmed by the *Planck* CIB power spectrum results compared to a similar modeling (Planck Collaboration XXX. 2013). For the bispectrum, this effect would also decrease the 1-halo term for the bispectrum, and most probably also the 2-halo term, compared to the 3-halo term. These changes would be consistent with a steeper-than-prescription bispectrum : e.g., the 3-halo term goes as $P(k)^2$ in equilateral which falls more steeply than the prescription going as $P(k)^{3/2}$.

To conclude, these CIB NG measurements open an interesting window for CIB science. They provide the opportunity to constrain CIB models with both 2-point and 3-point correlations, and they may indicate that the IR emissivity models need to depend not only on the redshift but also on other explicative parameters.

5.3 Measuring the thermal Sunyaev-Zel'dovich NG

As mentioned in Sect.4.2, the thermal Sunyaev-Zel'dovich (tSZ) effect is an important secondary anisotropy of the CMB, leaving a negative signal below 217 GHz and positive above⁶. As it probes preferentially the most massive galaxy clusters at low redshift ($z=0-1$), the tSZ signal is highly non-Gaussian. This section describes the tSZ bispectrum measurement that I have performed, in particular in the *Planck* context.

5.3.1 On simulations

Although the tSZ signal is highly non-Gaussian, its contribution to the NG in CMB maps is in general relatively lower than the other foregrounds (radio sources, CIB). For illustration, I have computed the bias that the tSZ introduces to local-type primordial NG estimation with Sehgal et al. (2010) simulations, and found that $|\Delta f_{\text{NL}}| < 1.5$ at all frequencies (see Table 5.9).

ν (GHz)	30	90	148	219	277	350
$\ell_{\text{max}} = 700$	-0.34	-0.18	-0.04	$5 \cdot 10^{-7}$	0.04	0.47
$\ell_{\text{max}} = 2048$	-0.92	-0.51	-0.11	$1 \cdot 10^{-6}$	0.11	1.30

Table 5.9: Bias on the primordial NG parameter f_{NL} due to the tSZ signal simulated by Sehgal et al. (2010).

⁶for relativistic electrons the zero frequency increases slightly with temperature

I have also computed the tSZ bispectrum on a simulation by the *Planck* Sky model (PSM, Delabrouille et al. 2013), with a binning $\Delta\ell = 64$ and $\ell_{\max} = 2048$. The Sehgal et al. (2010) simulations yield the same behaviour.

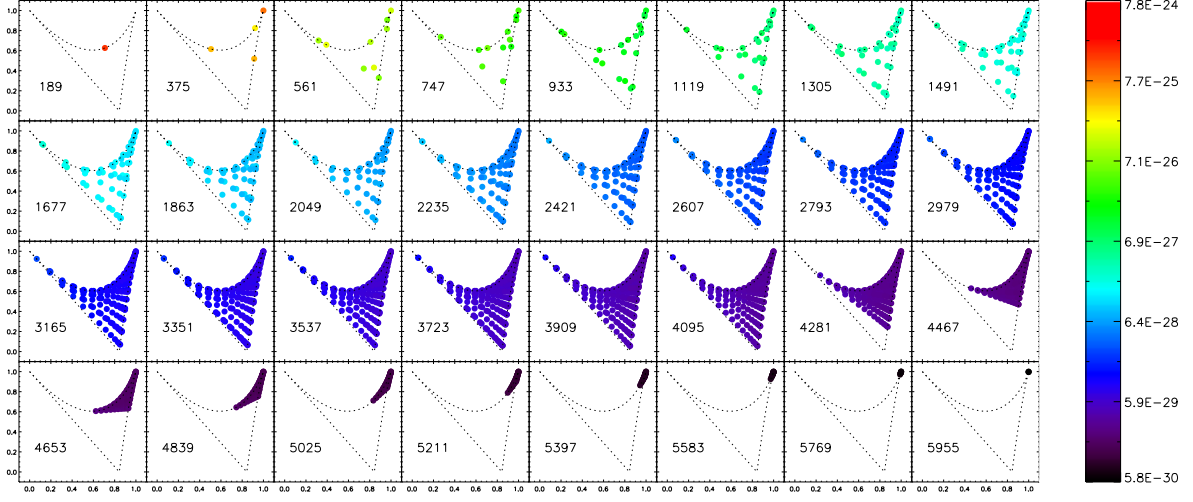


Figure 5.11: tSZ bispectrum computed from a simulation by the *Planck* Sky Model, plotted in the geometrical parametrisation.

We see from Fig.5.11 that the tSZ bispectrum is scale dependent but has little dependence on the shape of the triangle, as its value is nearly constant in each subplot. By analogy with the CIB, this means that the 1-halo term dominates the bispectrum (see e.g. Fig.4.14).

We see that the tSZ signal has a particular NG signature in comparison with the other signals studied previously. Indeed, the tSZ bispectrum depends on scale contrary to the radio bispectrum, and contrary to the CIB bispectrum it does not peak in squeezed and has little/no dependence on configuration. Non-Gaussianity is thus potentially a powerful tool to discriminate these foregrounds, in particular the tSZ and CIB whose power spectra have similar slopes on the range of multipoles of interest for *Planck* (Planck Collaboration XXI 2013).

5.3.2 *Planck* tSZ data : characteristics and processing pipeline

Due to its unique spectral signature, the tSZ signal can be detected with adapted filters. The large frequency range of *Planck* yields a data set of high quality for this effort. For example, the tSZ effect of about a thousand clusters has been revealed by *Planck* through adapted detection algorithms (Herranz et al. 2002; Melin et al. 2006; Carvalho et al. 2012). This sample and its characterisation form the *Planck* SZ catalogue (Planck Collaboration XXIX 2013).

Another project of the *Planck* Collaboration is the estimation of an all-sky tSZ map, mapping the Compton parameter y . To this end, numerous component separation methods have been proposed. I have participated in the assessment of their relative quality, the quality criterion being the conservation of the tSZ NG. Namely, the PSM produced maps of the different sky components, which were combined to yield *Planck* simulated frequency maps. Then the different component separation methods took these frequency

maps as input and estimated a tSZ map. Different tests were performed : recovery of the power spectrum, recovery of the fluxes and profile of the input clusters, and preservation of the NG character. I took charge of comparing the non-Gaussianity performance of the different component separation methods. I measured the bispectrum of the map produced by each method, with a galactic mask and a binning $\Delta\ell = 64$ and $\ell_{\max} = 2048$, and compared it with the bispectrum of the original tSZ map. Figure 5.12 shows the bispectrum of the full-sky original tSZ map (black), of the masked tSZ map debiased by f_{SKY} (red), and of the different component-separated maps debiased by f_{SKY} (colors).

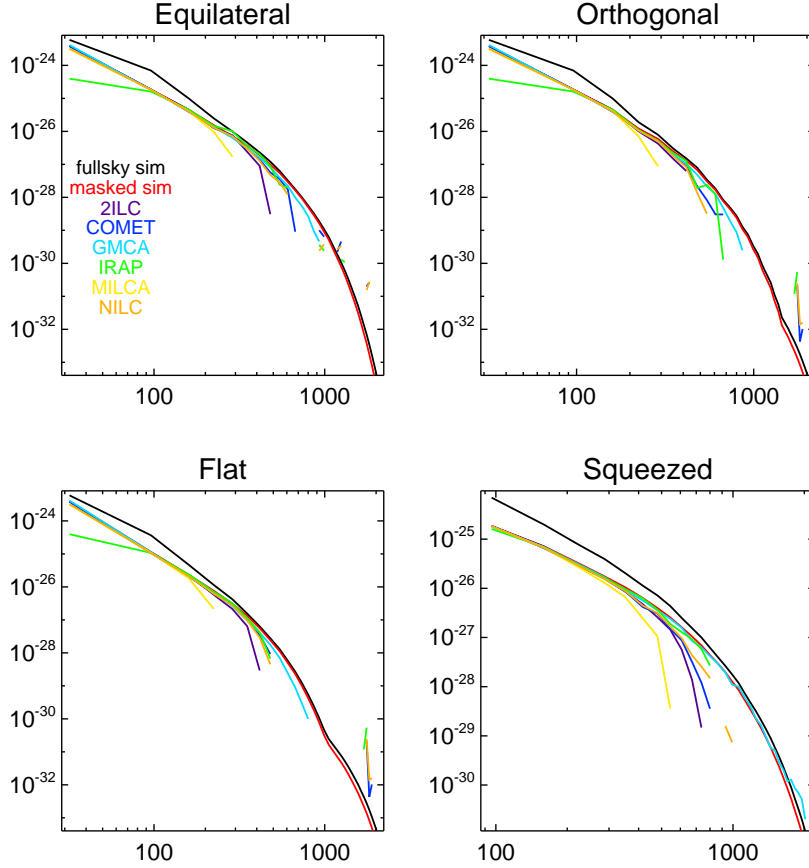


Figure 5.12: tSZ bispectrum from the input tSZ map (black full-sky, red masked) and its measurement on tSZ reconstructions by several component separation methods with the PSM simulations (Delabrouille et al. 2013), plotted in some chosen configurations.

We see that the methods have different performances. Most of them allow a good recovery of the bispectrum at low multipoles, although the IRAP method underestimates the first bin. At higher multipoles the methods tend to underestimate the bispectrum, then the line breaks when the bispectrum becomes negative. It afterwards becomes noise-dominated, alternating between positive and negative values. This break happens between $\ell = 250$ and $\ell = 900$ depending on the method, with NILC (Needlet Internal Linear Combination, Remazeilles et al. 2011) and GMCA (General Morphological Component Analysis, Bobin et al. 2013) performing the best, as they extend to higher multipoles.

It is worth noting that radio-sources are known to leave negative residuals in the MILCA maps (Modified Internal Linear Combination Algorithm, Hurier et al. 2013), and hence a negative bispectrum which should explain the break at high multipoles in Fig.5.12.

Finally, based on the different tests, the component separation methods that were selected for the analysis of *Planck* data were NILC and MILCA.

The *Planck* y-map article (Planck Collaboration XXI 2013) presents the first estimated full-sky maps of the tSZ signal. It uses the *Planck* data as well as the FFP6 (Full Focal Plane and version 6) simulations. These simulations use the PSM as signal model, and take it through the *Planck* data processing pipeline, from time-ordered input to the mapmaking stage, adding noise and various instrumental artefacts along the way, finally producing *Planck*-like frequency maps. Planck Collaboration XXI (2013) further quantifies the contamination of the recovered tSZ maps by foregrounds and assesses their consistency with the PCCS tSZ catalogue by comparing the detection of sources in the maps and the corresponding tSZ flux. Then the maps are analysed with the following statistics : angular power spectrum and bispectrum, and 1D p.d.f.. The power spectrum is estimated through cross-correlation of the first and last rings maps, and fitted with a halo model description with 1-halo and 2-halo terms (see Sect.4.2.2). As the tSZ power spectrum depends on the cosmology, the measurement allows to infer constraints on the cosmological parameters. Namely, the main cosmological parameters having an effect on C_ℓ^{tSZ} are σ_8 and Ω_m . The tSZ power spectrum analysis found the following constraint on these parameters : $\sigma_8 = 0.74 \pm 0.06$ and $\Omega_m = 0.33 \pm 0.06$, with a strong degeneracy along $\sigma_8 \cdot \Omega_m^{0.39} = \text{cste}$. The 1D p.d.f. analysis compared the recovered tSZ maps skewness to that of the FFP6 simulation and assumed that this skewness scaled as $\sigma_8^{10.7-11.1}$, it found the constraint $\sigma_8 = 0.779 \pm 0.015$. In the following, I describe the bispectrum analysis of the recovered tSZ maps.

As already shown in Fig.2.1, a Galactic and point-source mask was used for the tSZ analysis, leaving $\sim 60\%$ of the sky available. In order to reduce the variance of the bispectrum estimation, I measured the linear correction term discussed in Sect.2.2.4.3. This correction is negligible in most configurations except in squeezed, where it reaches 3.5% ⁷.

The other problem induced by the mask is the bias, see Sect.2.2.4.4. To correct for this, I have runned simulations using the method described in Sect.2.3.2. These simulations had a tSZ-like bispectrum with a high SNR, and I measured the ratio of the masked bispectrum to the full-sky bispectrum. The results are shown in Fig.5.13 and compared to the f_{SKY} approximation.

We see that contrary to the CIB (Fig.5.7), the ratio does not follow the f_{SKY} approximation. Instead, compared to the f_{SKY} approximation, the ratio shows a deficit at low multipoles and an excess in the squeezed configurations at high multipoles ($\ell \sim 1000$). The ratio becomes negative in equilateral and orthogonal for some high multipoles configurations. This situation is due to the fact that the tSZ bispectrum is much steeper than the CIB one, so that the large scale power leaks onto the small scales. The negative

⁷The smallness of the linear correction is due to the fact that the multipole bins are large, while the correction peaks in squeezed because the mask induces mostly a leakage of large scales power to small scales

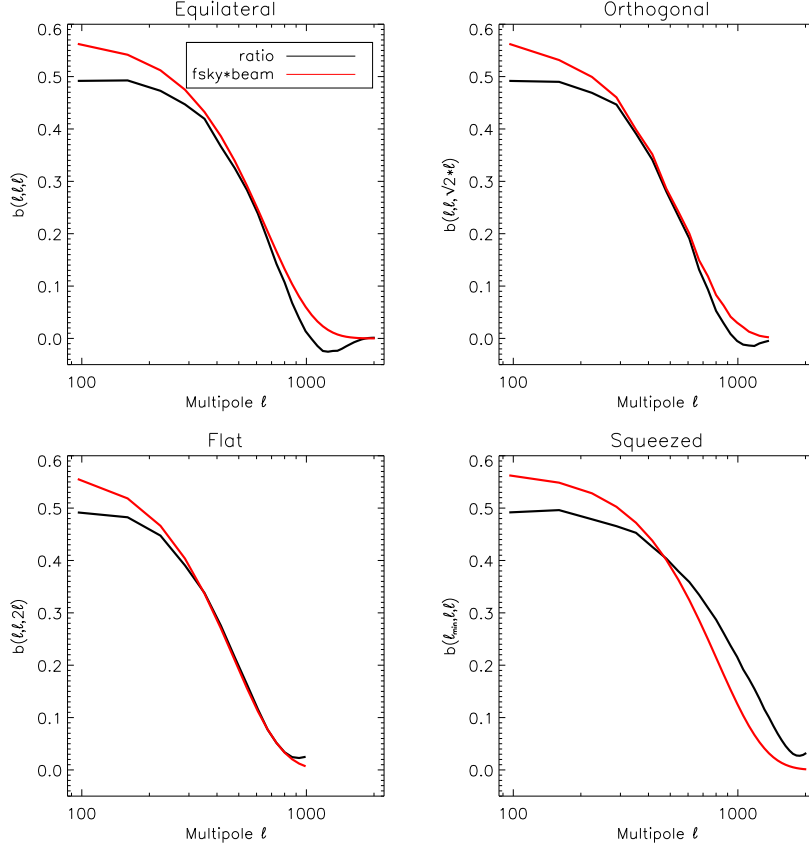


Figure 5.13: Masked sky to full-sky bispectrum ratio with simulations of a tSZ-like bispectrum, compared with the f_{SKY} approximation.

values of the ratio can be interpreted as resulting from the point-source mask. Indeed, the latter yields numerous small areas with zero value (while the y -map is positive) yielding negative skewness on small angular scales.

When the measured ratio is too different from the f_{SKY} approximation, the mask-induced leakage is too important and we cannot uncover the tSZ bispectrum. This happens in particular at high multipoles, where the tSZ bispectrum estimates indeed diverge. In the analysis, I have discarded affected configurations, for which the measured ratio differs from the f_{SKY} approximation by more than 50%. This excluded most configurations with a multipole higher than 1024, and I could run new simulations at $\ell_{\text{max}} = 1024$ and use a resolution of $N_{\text{side}} = 512$. I ran 100 simulations at $N_{\text{side}} = 512$. They confirmed Fig.5.13, and showed that the ratio converged.

Having taken care of the mask effect, we also need an in-depth assessment of the foreground effects, i.e., of their possible contamination to the bispectra. To this end, the FFP6 (Full Focal Plane and version 6) simulations were used. These simulations use the PSM as signal model, and take it through the *Planck* data processing pipeline, from time-ordered input to the mapmaking stage, adding noise and various instrumental artefacts along the way, finally producing *Planck*-like frequency maps. NILC and MILCA methods use the frequency maps as input and estimate a y -map. To test the foreground

contamination, NILC/MILCA weights were applied to the original foregrounds maps to build residuals maps (since the methods are linear). I have computed the bispectrum of the residuals maps and of the tSZ map, and Fig.5.14 shows these results in the case of the MILCA method (the absolute value of the bispectrum is shown for the residuals).

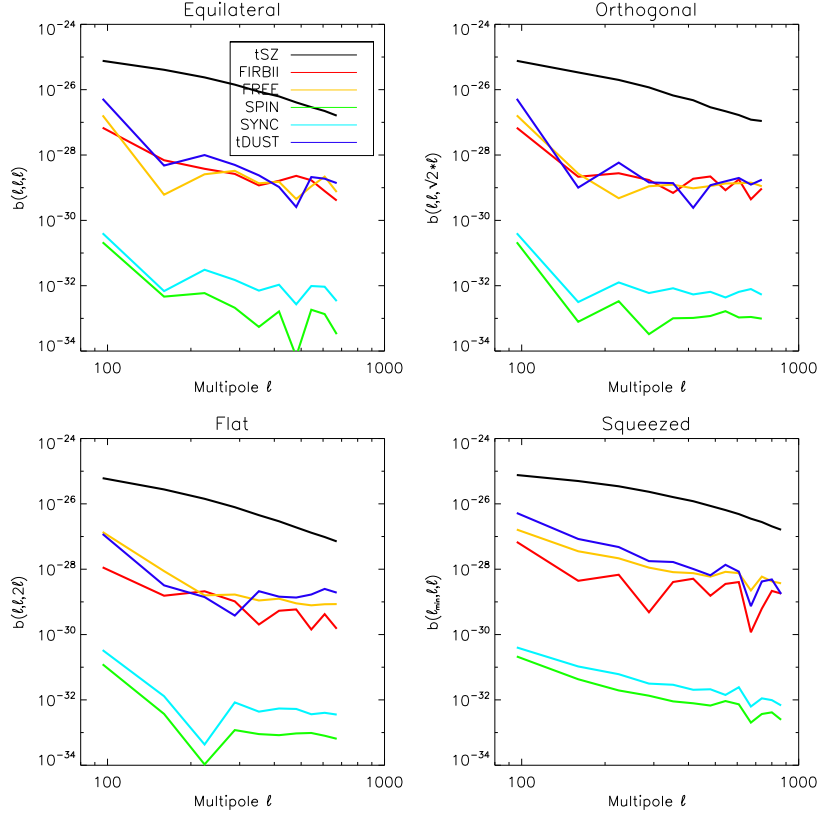


Figure 5.14: From Planck Collaboration XXI (2013), comparison of the tSZ bispectrum and the foregrounds residuals (in absolute value) for the MILCA (Hurier et al. 2013) component separation method on FFP6 simulations.

The configurations involving the first multipole bin are too contaminated by foregrounds ; these configurations are already discarded in Fig.5.14 as a result of this observation. For the other configurations, Fig.5.14 shows that the bispectrum of foregrounds residuals is much smaller than the tSZ bispectrum. Hence the bispectrum estimates of the recovered y -maps are safe from foreground contamination in the considered configurations.

Then, I have discarded the configuration which fulfilled any of the following conditions :

- masked ratio differing from the f_{SKY} approximation by more than 50%, or negative
- one bin is the first one, and is too contaminated by foregrounds
- one of the multipole is greater than 1000

- the NILC or MILCA estimate is lower than the minimum acceptable value 10^{-29} ⁸

where the third and fourth conditions were put in place to discard some remaining configurations which had escaped the first two tests but were nonetheless unreliable.

5.3.3 *Planck* tSZ data : results and characterisation

Considering all the corrections and flagging described in the previous section, Fig.5.15 shows my final estimates of the tSZ bispectrum for the NILC and MILCA recovered y -maps with their error bars.

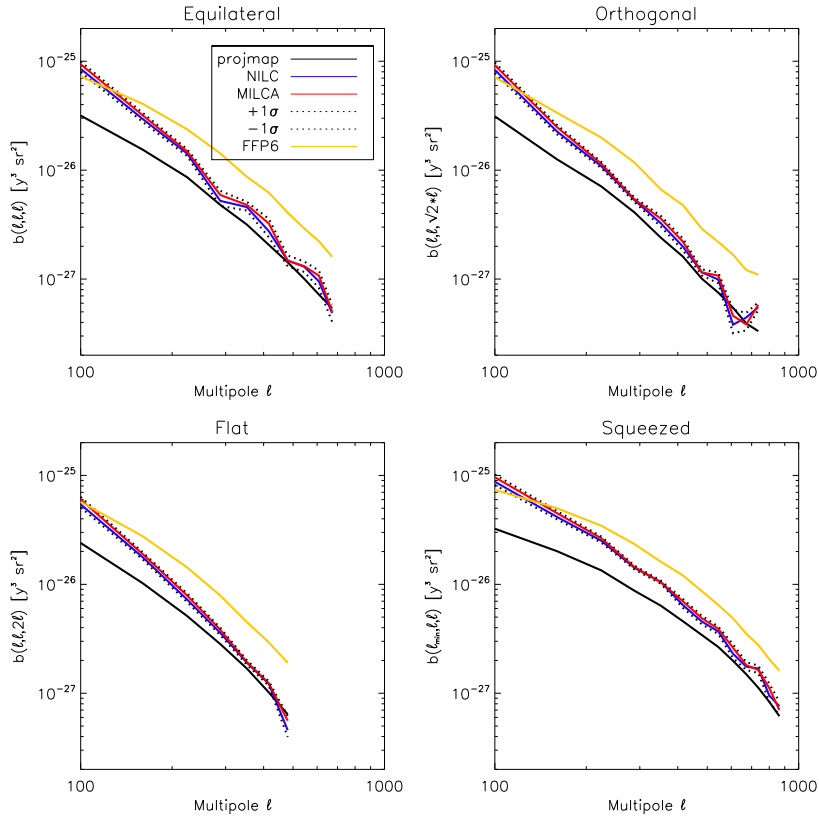


Figure 5.15: From Planck Collaboration XXI (2013), tSZ bispectrum estimates with NILC and MILCA recovered y -maps, with their error bars in dotted lines. Overplotted are the bispectrum of the projected catalogue from Planck Collaboration XXIX (2013) and of the FFP6 simulation.

We see that the bispectrum estimates for the NILC and MILCA y -maps are consistent with each other. The estimated bispectrum for NILC is lower on average, and they have the same order of magnitude as the FFP6 tSZ simulation, although being steeper. I have also overplotted (in black) an estimation of the tSZ bispectrum from the detected clusters present in the *Planck* SZ catalogue (Planck Collaboration XXIX 2013). This estimate was computed with a y -map provided the detected (and confirmed) clusters were

⁸This minimum acceptable value is ~ 2 orders of magnitude below the minimum measured value in Fig.5.15.

projected with the tSZ flux and radius from the catalogue. We see that the bispectrum estimates from NILC and MILCA have the same order of magnitude as the projected map, but are significantly larger. This indicates a contribution from undetected clusters or diffuse tSZ emission, or possibly foreground contamination at low multipoles.

The $\pm 1\sigma$ error bars around the bispectrum estimates from NILCA and MILCA y -maps are shown with dotted lines, they were computed with the following terms :

- cosmic variance : $C_{2\times 2\times 2}$ term with the f_{SKY} approximation. For the power spectrum involved, I estimated the NILC and MILCA y -map auto-spectra (containing the noise power spectrum, as necessary), and averaged them.
- cosmic variance : $C_{3\times 3}$ term with the f_{SKY} approximation. The bispectrum involved is the estimated average of the bispectra from NILC and MILCA y -maps.
- cosmic variance : the $C_{4\times 2}$ and C_6 terms were neglected, as I did not have estimates of the tSZ trispectrum nor 6-point function.
- systematic effects : to quantify the systematic uncertainty due to the component separation method, I used the difference between the bispectra from NILC and MILCA y -maps.

Note that the $C_{3\times 3}$ term produces a non-diagonal covariance matrix, hence the error bars in Fig.5.15 are correlated.

The significance of the tSZ NG detection is quantified by computing the SNR in different cases summarised in Table 5.10.

SNR	NILC	MILCA
full non-Gaussian	210.61	223.29
Gaussian	332.49	351.62
non-Gaussian diagonal	319.77	338.47

Table 5.10: Detection SNR for the tSZ NILC and MILCA bispectra, in different cases. The ‘Gaussian’ case neglects the $C_{3\times 3}$ term, while the ‘non-Gaussian diagonal’ only considers its contribution to the variance, so that the covariance matrix is diagonal.

The first line (‘full non-Gaussian’) shows the SNR with the full covariance matrix described above. We see that the result are highly significant. There are 416 non-flagged configurations, each being on average a 14.2σ (15.3σ) detection for the NILC (MILCA) bispectrum. This can be compared favorably with the $\sim 10\sigma$ announced detection of the tSZ bispectrum by Crawford et al. (2013), who consider all configurations and all SPT frequencies.

Furthermore, the Crawford et al. (2013) analysis only accounts for Gaussian contributions to the error bars. For comparison, I computed the SNR that would be obtained by dismissing the $C_{3\times 3}$ term of the covariance matrix, it is shown in the first line of Table 5.10. I also computed the SNR that would be obtained considering $C_{3\times 3}$ contribution to the bispectrum variance but not the off-diagonal contributions, it is shown in the second line of the same Table. We see that considering the full non-Gaussian covariance matrix is important, since we otherwise overestimate the detection significance by

$\sim 60\%$. We also see that the decrease of SNR mostly comes from the off-diagonal terms of the covariance matrix. The non-Gaussian term may thus seem small when considering the variance, but it is important to consider them for the whole significance computation.

We can now consider the configuration dependence of the tSZ bispectrum measurements. To this end, I have plotted the bispectra for NILC and MILCA y -maps with the geometrical parametrisation. For illustration, the MILCA bispectrum is shown in Fig.5.16⁹.

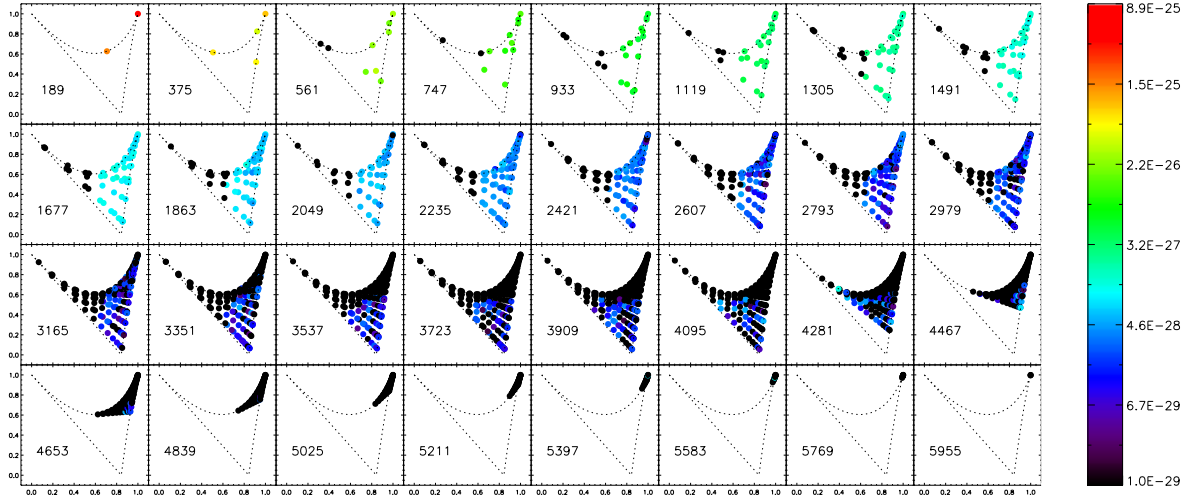


Figure 5.16: tSZ bispectrum estimate for MILCA y -map, plotted in the geometrical parametrisation. Black dots denote flagged configurations.

This figure shows a behaviour remarkably consistent with Fig.5.11. The bispectrum varies with scale but does not depend on the shape of the triangle. This gives further confidence in the robustness of the measurement, since the bispectra of other foregrounds do not exhibit this behaviour. We also note the heavy flagging of configurations, especially for squeezed triangles and at high multipoles.

Finally, I used these non-Gaussianity measurements for cosmological inference. Indeed, Bhattacharya et al. (2012) published a theoretical study of the tSZ bispectrum and showed its dependence with cosmological parameters. Specifically, considering the most influential parameters, they showed that the amplitude of the bispectrum scales as $\sigma_8^{11.6} \Omega_b^{4.1}$. I compared the amplitude of the measured bispectrum with that of the FFP6 simulation (whose cosmological parameters are known) in the equilateral configurations. I could thus put a constrain on these parameters (σ_8, Ω_b). Considering the measurement error bars, the dispersion between NILC and MILCA, and an uncertainty on the σ_8 exponent¹⁰, I have the conservative constrain (Planck Collaboration XXI 2013) :

$$\sigma_8 \left(\frac{\Omega_b}{0.049} \right)^{0.35} = 0.74 \pm 0.04 \quad (5.36)$$

⁹The NILC bispectrum shows the same behaviour.

¹⁰Bhattacharya et al. (2012) give an uncertainty on the σ_8 exponent, being between 11 and 12. I was conservative and took a range [10,13] for this exponent.

It is remarkably consistent with the other constraints on σ_8 from tSZ measurements, either from the power spectrum or the skewness (Planck Collaboration XXI 2013), or from the catalogue of detected clusters (Planck Collaboration XX 2013).

I have thus provided the first measurement of the tSZ bispectrum by configuration. The measurement over two different y -map estimation methods are found consistent together, and the shape of the bispectrum is found consistent with expectations from simulations. Furthermore, the excellent signal-to-noise ratio of the detection opens up the possibility of using tSZ non-Gaussianity to constrain tSZ physics and cosmology, and I indeed realised first constraints on the cosmological parameters σ_8 and Ω_b .

A full inference analysis with a tSZ model would however be necessary to fully exploit these bispectrum measurements, and it is indeed one of my next goals to build such a model and the related inference analysis.

Conclusion

In this thesis, I have shown how random fields can be characterised by their (connected) correlation functions. Especially, in harmonic space these correlation functions are called polyspectra, and I have shown their properties on the sphere. I have shown how they can be estimated at second and third order (power spectrum and bispectrum), even in the presence of incomplete sky coverage. For the bispectrum case, the incomplete sky coverage is computationally problematic as the coupling matrix cannot be computed numerically from its analytical definition. I have designed a method to compute this matrix through simulations, but it was still unmanageable numerically for a *Planck*-like resolution. Thus, when analysing *Planck* data I have introduced an approximate method to debias the bispectrum estimate and locate cases where the approximation fails.

For cosmology, the study of high order correlation functions is of interest for the CMB, in particular as a possibility to discriminate the primordial process generating the cosmological perturbation (e.g. inflation models). As the foregrounds to the CMB have a non-Gaussian distribution, they may bias/contaminate the study of the primordial NG. My study has focused on the extragalactic foregrounds, and first on radio point-sources and the Cosmic Infrared Background. I have proposed in Lacasa et al. (2012) an empirical and analytical prescription for the non-Gaussianity of clustered point-sources accounting for possible multiple populations. Based on this prescription, I built in Lacasa & Aghanim (2013) a fast estimator for the amplitude of the CIB bispectrum. I further showed how to combine this estimator with corresponding estimates of the radio and primordial NG, in order to build a joint and robust estimation of non-Gaussianity. It has allowed me in particular to place upper limits on the contamination of point-sources to the primordial NG estimation (Planck Collaboration XXIV 2013).

I have also developed tools necessary for physical modeling of high orders for LSS tracers. I have indeed shown how 3D polyspectra can be projected onto the sphere, and developed a diagrammatic method to compute the polyspectra of the galaxy density field. I have shown the resulting equations for the bispectrum of the tSZ signal, and more thoroughly for the Cosmic Infrared Background. For the latter, I have implemented a numerical computation of the bispectrum, which allowed the first predictions of the CIB bispectrum and its variation with model parameters (Lacasa et al. 2013). I have shown the interest of this quantity, as the constraints on CIB model parameters were of similar quality compared to the constraint coming from the power spectrum (Pénin et al. 2013).

Finally, I have made the first measurement of the foregrounds NG on *Planck* data. I have computed the first measurement of the tSZ bispectrum, using tSZ maps estimated by *Planck*. This measurement proves consistent with expectations from the *Planck* tSZ catalogue and with available tSZ simulations; it is furthermore of high significance. Although I did not have a theoretical model, and its numerical implementation, to compare

to the measurement, I was able to use the measurement to put a constraint on cosmological parameters, specifically on a combination of σ_8 and Ω_b (Planck Collaboration XXI 2013).

I have also realised the first measurements of the CIB bispectrum at several frequencies, accounting for the problem of partial sky coverage and contamination by Galactic dust in particular. The detection is significant at each frequency, with significance increasing from 217 to 545 GHz, and the estimates are consistent together. Indeed the intrinsic level of non-Gaussianity as well as the slope of the bispectrum are found consistent between frequencies (Planck Collaboration XXX. 2013). The measured bispectrum is however found significantly steeper than the empirical prescription or the halo model prediction presented in this thesis, which hence calls for a complexification of the IR emissivity model.

Perspectives

There are several possible developments and improvements of the work that I have presented in this thesis. I will list below some possibilities, grouped by thematics.

First, the effects on non-Gaussianity studies of partial sky coverage need to be better accounted for. Indeed, the problem of the leakage of large scale power onto small scales has been limiting my estimates of the tSZ bispectrum with *Planck* data. It is also limiting the CIB bispectrum estimates because of the Galactic dust contamination. Coupling matrix computations are needed to quantify this effects, and the numerical method I have presented to compute this matrix becomes too CPU intensive at high resolution. However this method may be manageable with a better choice of bispectrum basis and/or new methods to generate non-Gaussian maps. We may look for a method to optimise the α_ℓ in my simulated bispectra, e.g. by orthogonalising the bispectra with the Gram-Schmidt process, or we may use cross-bispectra between different simulations. Alternatively other methods may be searched for, the separable modal methodology (Fergusson & Shellard 2007, 2009) may provide useful tools for example.

For the *Planck* data analysis presented in this thesis, a first possibility to mitigate the mask effect would be a high-pass filtering by estimating low multipoles through maximum likelihood and subtracting them. The robustness of the results with respect to this filtering should then be tested. This filtering may in particular mitigate the dust contamination to the CIB bispectrum estimates.

Using other bases, e.g. wavelets/needlets (e.g. Pietrobon et al. 2009), or other quantities, e.g. Minkowsky functionals (e.g. Ducout et al. 2013), is another possibility, as they may account more easily for the mask effect. However, theoretical predictions would need to be propagated to the new basis, which may not be straightforward. For example the wavelet-based f_{NL} estimator for *Planck* relied on simulations to compute the theoretical third order wavelet statistics produced by primordial NG. The modal methodology has also emerged as a powerful possibility, allowing a fast near-complete and model-independent description of the NG at third order. It may be even more powerful at higher orders compared to current methods, although physical understanding of the results and control of the systematics may be more difficult in this case.

Then, I would like to develop a model for the tSZ bispectrum with its numerical implementation, as I did for the CIB bispectrum. This model is needed to exploit the information present in the *Planck* tSZ bispectrum measurement. In particular we could combine the information coming from the power spectrum measurement with that of the bispectrum, increasing the constraint on model parameters. More generally, I am interested in building a more optimal analysis of the tSZ signal, which is highly non-Gaussian. This may either be done through non-linear data transformation (Leclercq et al. 2013; Carron & Szapudi 2013; Simpson et al. 2013), or Bayesian analysis may be

possible if we neglect halo correlations.

The tSZ bispectrum model would also serve to explore physically motivated additions to the Λ CDM cosmology. For example, massive neutrinos (Mak & Pierpaoli 2013) and primordial NG (Mak & Pierpaoli 2012; Trindade et al. 2013) have been shown to have a significant impact on the tSZ signal, while the halo triaxiality (Limousin et al. 2013) has an impact on the galaxy bispectrum (Smith et al. 2006) and thus most probably also on the tSZ bispectrum.

Furthermore, the model for the CIB bispectrum I have presented, based on the halo model, need to be complexified to account for the *Planck* measurements. As already discussed in Sect.5.2.3 a promising track to reconcile model and data is to consider new IR emissivities, e.g. depending on the halo mass and/or on the galaxy type (central/satellite) such as in the model by Béthermin et al. (2013) and Planck Collaboration XXX. (2013). This more complex modeling can be incorporated in the formalism described in Sect.4.3.3, and in particular the diagrammatic method can be adapted straightforwardly. We may also want to assess the effect of halo triaxiality which has a noticeable effect on the galaxy bispectrum as already mentioned.

Besides, the cross-correlation between the CIB and tSZ effect would be interesting to explore. Indeed, their cross-spectrum has been predicted by Addison et al. (2012), and was needed to debias the *Planck* power spectrum estimates both for the tSZ study and for the CIB study. The Addison et al. (2012) model revealed unpractical to use within *Planck*, and for a consistent treatment the modeling of the tSZ and CIB signals must be the same as in respectively Planck Collaboration XXI (2013) and Planck Collaboration XXX. (2013). Moreover a cross-spectrum is not the only quantity of interest scientifically. For example, one would want to assess the CIB contamination to the estimated tSZ flux in detected clusters. Also, as both signals are non-Gaussian, a cross-spectrum does not allow for an optimal analysis of the cross-information between the signals.

Moreover, CMB and galaxy lensing are research areas where the tools I have developed could be adapted. Indeed, the study of these signals involves the modeling of higher orders of the distribution of the gravitational potential sourced by the large-scale structure of the universe. Thus non-Gaussianity tools as well as the halo model at high orders are of particular interest in this prospect.

Finally, there are several foregrounds or secondary anisotropies tracing the LSS : CMB lensing, iSW, tSZ, CIB, and surveys in other wavelengths (radio, galaxy surveys in the optical and near-IR, cluster in X-rays...). Most of their cross-spectra have been or are currently been studied, however the question remains open as of how to best combine these probes. In this prospect, I am thus also interested in looking for statistical methods to extract more information from a joint analysis of these signals.

Glossary

Symbol	Definition
\hat{A}	estimator of a quantity A
A_{IR}	CIB bispectrum amplitude
$a_i, a_{\ell_i m_i}$	harmonic coefficient
B_ℓ	beam profile in harmonic space
$b_{123}, b_{\ell_{123}}, b_{\ell_1 \ell_2 \ell_3}$	angular bispectrum
\mathbf{b}_{123}	bispectrum accounting for the beam effect
b_{PS}	radio sources bispectrum amplitude
$C_{2 \times 2 \times 2}, C_{3 \times 3}, C_{4 \times 2}, C_6$	terms of the bispectrum cosmic covariance
\mathcal{C}_ℓ	power spectrum accounting for the beam and noise effect
$F_{\alpha\beta}$	Fisher matrix element
f_{NL}	primordial NG amplitude parameter
F^s	second order perturbation theory kernel
f_{SKY}	fraction of sky
$G_{123}, G_{\ell_1 \ell_2 \ell_3}^{m_1 m_2 m_3}$	Gaunt coefficient
g_ν	tSZ frequency dependence
\vec{k}	2D Fourier mode
\mathbf{k}	3D Fourier mode
\hat{n}	direction, a unit vector
n_{config}	number of bispectrum configurations
$N_{\ell_1 \ell_2 \ell_3}$ or N_{123}	normalisation of the bispectrum estimator, \sim number of triangles
$\mathcal{P}^{(n)}$	n -th order polyspectrum
$r(z)$	comoving distance
$Y_i, Y_{\ell_i m_i}$	spherical harmonic function
$\Delta_{\ell_1 \ell_2 \ell_3}$	$\begin{cases} 6 & \text{equilateral triangles} \\ 2 & \text{isosceles triangles} \\ 1 & \text{general triangles} \end{cases}$
κ_n	cumulant of order n
μ_n	centered moment of order n
ν	frequency
σ_8	amplitude parameter for the density perturbations
Ω_i	normalised energy density of fluid i
Ω_{pix}	pixel surface area

Abbreviation	Signification
ACT	Atacama Cosmology Telescope
AGN	Active Galactic Nucleus
AME	Anomalous Microwave Emission
CDM	cold dark matter
c.f.	correlation function
c.g.f.	cumulant generating function
CIB	Cosmic Infrared Background
CMB	Cosmic Microwave Background
DM	dark matter
d.o.f.	degree(s) of freedom
DSFG	dusty star-forming galaxy
ERCSC	Early Release Compact Source Catalog (2011)
FFP6	Full Focal Plane version 6 (simulations)
FL	Friedman-Lemaitre
FWHM	Full Width to Half Maximum
GMCA	General Morphological Component Analysis
GR	General Relativity
GUT	Grand Unification Theory
HFI	High Frequency Instrument
HOD	Halo Occupation Distribution
iff	if and only if
i.i.d.	independent identically distributed
IR	infrared
IRAS	Infrared Astronomical Satellite
ISW	integrated Sachs-Wolfe
KSW	Komatsu-Spergel-Wandelt (estimator)
kSZ	kinetic Sunyaev-Zel'dovich (effect)
LFI	Low Frequency Instrument
LSS	Large Scale Structure
m.g.f.	moment generating function
MILCA	Modified Internal Linear Combination Algorithm
NFW	Navarro-Frenk-White (profile)
NILC	Needlet Internal Linear Combination
NG	non-Gaussianity
PCCS	Planck Catalogue of Compact Sources (2013)
p.d.f.	probability density function
PNG	primordial non-Gaussianity
ps or p.s.	point-source(s)
PSM	<i>Planck</i> Sky Model
rms	root mean squared
SNR	signal-to-noise ratio
SPT	South Pole Telescope
SW	Sachs-Wolfe (effect)
SZ	Sunyaev-Zel'dovich (effect)
tSZ	thermal Sunyaev-Zel'dovich (effect)
WMAP	Wilkinson Microwave Anisotropy Probe

Bibliography

- Abramowitz M., Stegun I.A., 1964, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover Publications
- Addison G.E., Dunkley J., Spergel D.N., Dec. 2012, MNRAS, 427, 1741
- Aghanim N., Desert F.X., Puget J.L., Gispert R., Jul. 1996, A&A, 311, 1
- Aghanim N., Majumdar S., Silk J., Jun. 2008, Reports on Progress in Physics, 71, 066902
- Amblard A., Cooray A., Serra P., et al., Feb. 2011, Nature, 470, 510
- Argüeso F., González-Nuevo J., Toffolatti L., Nov. 2003, ApJ, 598, 86
- Austermann J.E., Dunlop J.S., Perera T.A., et al., Jan. 2010, MNRAS, 401, 160
- Babich D., Aug. 2005, Physical Review D, 72, 043003
- Babich D., Pierpaoli E., Jun. 2008, Phys. Rev. D, 77, 123011
- Bartlett J.G., Feb. 2004, Astrophysics and Space Science, 290, 105
- Bartlett J.G., Amram P., Apr. 1998, ArXiv Astrophysics e-prints
- Bartolo N., Komatsu E., Matarrese S., Riotto A., Nov. 2004, Physics Reports, 402, 103
- Bayes T., Price R., 1763, Philosophical Transactions of the Royal Society of London, 53, 370
- Bennett C.L., Halpern M., Hinshaw G., et al., Dec. 1997, In: American Astronomical Society Meeting Abstracts, vol. 29 of Bulletin of the American Astronomical Society, 1353
- Berlind A.A., Weinberg D.H., Aug. 2002, ApJ, 575, 587
- Bernardeau F., 2007, Cosmologie, des fondements théoriques aux observations, CNRS Editions, EDP Sciences
- Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., Sep. 2002, Phys. Rep., 367, 1
- Bernardeau F., Pitrou C., Uzan J.P., Feb. 2011, J. Cosmology Astropart. Phys., 2, 15
- Béthermin M., Dole H., Feb. 2011, ArXiv e-prints

- B  thermin M., Wang L., Dor   O., et al., Apr. 2013, ArXiv e-prints
- Bhattacharya S., Nagai D., Shaw L., Crawford T., Holder G.P., Nov. 2012, The Astrophysical Journal, 760, 5
- Blas D., Lesgourgues J., Tram T., Jul. 2011, J. Cosmology Astropart. Phys., 7, 34
- Bobin J., Starck J.L., Sureau F., Basak S., Feb. 2013, A&A, 550, A73
- Boggess N.W., Mather J.C., Weiss R., et al., Oct. 1992, ApJ, 397, 420
- Boothroyd A.I., Blagrove K., Lockman F.J., et al., Dec. 2011, A&A, 536, A81
- Boulanger F., Abergel A., Bernard J.P., et al., Aug. 1996, A&A, 312, 256
- Brown I.A., May 2013, ArXiv e-prints
- Brown R.G., 2009, Classical Electrodynamics, Robert Brown, URL <http://www.phy.duke.edu/~rgb/Class/phy319/phy319/phy319.html>
- Bucher M., Tent B.V., Carvalho C.S., Oct. 2010, Monthly Notices of the Royal Astronomical Society, 407, 2193
- Buchert T., Jan. 1993, A&A, 267, L51
- Carron J., Sep. 2011, The Astrophysical Journal, 738, 86
- Carron J., Szapudi I., Jun. 2013, ArXiv e-prints
- Carvalho P., Rocha G., Hobson M.P., Lasenby A., Dec. 2012, MNRAS, 427, 1384
- Condon E.H., Shortley G.H., 1935, The Theory of Atomic Spectra, Cambridge University Press
- Cooray A., Sheth R., Dec. 2002, Phys. Rep., 372, 1
- Cooray A., Baumann D., Sigurdson K., 2005, In: Melchiorri F., Rephaeli Y. (eds.) Background Microwave Radiation and Intracluster Cosmology, 309
- Coppin K., Chapin E.L., Mortier A.M.J., et al., Nov. 2006, MNRAS, 372, 1621
- Crawford T.M., Schaffer K.K., Bhattacharya S., et al., Mar. 2013, ArXiv e-prints
- Creminelli P., Zaldarriaga M., Oct. 2004, Journal of Cosmology and Astroparticle Physics, 2004, 006
- Creminelli P., Nicolis A., Senatore L., Tegmark M., Zaldarriaga M., May 2006, Journal of Cosmology and Astroparticle Physics, 2006, 004
- Curto A., Mart  nez-Gonz  lez E., Barreiro R.B., Hobson M.P., Oct. 2011, MNRAS, 417, 488
- Dalal N., Dor   O., Huterer D., Shirokov A., Jun. 2008, Phys. Rev. D, 77, 123514

- Davies R.D., Wilkinson A., Apr. 1998, ArXiv Astrophysics e-prints
- Delabrouille J., Cardoso J., Feb. 2007, ArXiv Astrophysics e-prints
- Delabrouille J., Betoule M., Melin J.B., et al., May 2013, A&A, 553, A96
- Dicke R.H., Peebles P.J.E., Roll P.G., Wilkinson D.T., Jul. 1965, ApJ, 142, 414
- Dolney D., Jain B., Takada M., Aug. 2004, MNRAS, 352, 1019
- Donzelli S., Hansen F.K., Liguori M., Marinucci D., Matarrese S., Aug. 2012, ApJ, 755, 19
- Draine B.T., Dec. 2004, In: American Astronomical Society Meeting Abstracts, vol. 36 of Bulletin of the American Astronomical Society, 162.01
- Ducout A., Bouchet F.R., Colombi S., Pogosyan D., Prunet S., Mar. 2013, MNRAS, 429, 2104
- Dunne L., Eales S., Edmunds M., et al., Jun. 2000, MNRAS, 315, 115
- Efstathiou G., Apr. 2004, MNRAS, 349, 603
- Eisenstein D.J., Hu W., Jan. 1999, ApJ, 511, 5
- Fergusson J.R., Liguori M., 2010, arXiv:1006.1642, 1–30
- Fergusson J.R., Shellard E.P.S., Oct. 2007, Phys. Rev. D, 76, 083523
- Fergusson J.R., Shellard E.P.S., Aug. 2009, Phys. Rev. D, 80, 043510
- Fixsen D.J., Mather J.C., Dec. 2002, ApJ, 581, 817
- Fixsen D.J., Dwek E., Mather J.C., Bennett C.L., Shafer R.A., Nov. 1998, ApJ, 508, 123
- Fry J.N., Apr. 1984, The Astrophysical Journal, 279, 499
- Fry J.N., Gaztanaga E., Aug. 1993, ApJ, 413, 447
- Gamow G., Aug. 1948, Physical Review, 74, 505
- Giannantonio T., Jun. 2010, In: Alimi J.M., Fuözfa A. (eds.) American Institute of Physics Conference Series, vol. 1241 of American Institute of Physics Conference Series, 209–215
- Gispert R., Lagache G., Puget J.L., Aug. 2000, A&A, 360, 1
- González-Nuevo J., Toffolatti L., Argüeso F., Mar. 2005, ApJ, 621, 1
- Górski K.M., Hivon E., Banday A.J., et al., Apr. 2005, ApJ, 622, 759
- Guth A.H., Jan. 1981, Phys. Rev. D, 23, 347
- Guth A.H., Pi S.Y., Oct. 1982, Physical Review Letters, 49, 1110

- Hand N., Addison G.E., Aubourg E., et al., Jul. 2012, Physical Review Letters, 109, 041101
- Hanson D., Smith K.M., Challinor A., Liguori M., Oct. 2009, Phys. Rev. D, 80, 083004
- Hanson D., Challinor A., Lewis A., Sep. 2010, General Relativity and Gravitation, 42, 2197
- Hawking S.W., Sep. 1982, Physics Letters B, 115, 295
- Hazewinkel M., 2001, Encyclopedia of Mathematics, Springer
- Hermite C., 1864, In: Sur un nouveau développement en série de fonctions, vol. 58, 93–100 and 266–273, Compt. Rend. Acad. Sci. Paris
- Herranz D., Sanz J.L., Hobson M.P., et al., Nov. 2002, MNRAS, 336, 1057
- Hivon E., Górski K.M., Netterfield C.B., et al., Mar. 2002, ApJ, 567, 2
- Huang Z., Vernizzi F., Dec. 2012, ArXiv e-prints
- Hurier G., Macias-Perez J.F., Hildebrandt S.R., Jul. 2013, ArXiv e-prints
- Ilic S., 2013, The large scale structures : a window on the Dark components of the Univers, Ph.D. thesis, Université Paris-Sud
- Joachimi B., Shi X., Schneider P., Dec. 2009, A&A, 508, 1193
- Kaiser N., Sep. 1984, ApJ, 284, L9
- Kashlinsky A., Arendt R.G., Ashby M.L.N., et al., Jul. 2012, ApJ, 753, 63
- Kim J., Rotti A., Komatsu E., Apr. 2013, J. Cosmology Astropart. Phys., 4, 21
- Klypin A.A., Sazhin M.V., Strukov I.A., Skulachev D.P., Apr. 1987, Soviet Astronomy Letters, 13, 104
- Knox L., Cooray A., Eisenstein D., Haiman Z., Mar. 2001, ApJ, 550, 7
- Knox L., Holder G.P., Church S.E., Sep. 2004, ApJ, 612, 96
- Kogut A., 1999, In: de Oliveira-Costa A., Tegmark M. (eds.) Microwave Foregrounds, vol. 181 of Astronomical Society of the Pacific Conference Series, 91
- Komatsu E., Jun. 2002, ArXiv Astrophysics e-prints
- Komatsu E., Seljak U., Nov. 2002, MNRAS, 336, 1256
- Komatsu E., Wandelt B.D., Spergel D.N., Banday A.J., Górski K.M., Feb. 2002, ApJ, 566, 19
- Komatsu E., Kogut A., Nolte M.R., et al., Sep. 2003, Astrophysical Journal Supplement Series, 148, 119

- Komatsu E., Spergel D.N., Wandelt B.D., Nov. 2005, *The Astrophysical Journal*, 634, 14, URL <http://stacks.iop.org/0004-637X/634/i=1/a=14>
- Komatsu E., Smith K.M., Dunkley J., et al., Feb. 2011, *Astrophysical Journal Supplement Series*, 192, 18
- Kosowsky A., Dec. 2003, *New Astronomy Reviews*, 47, 939
- Kravtsov A.V., Berlind A.A., Wechsler R.H., et al., Jul. 2004, *ApJ*, 609, 35
- Lacasa F., Aghanim N., Nov. 2013, *ArXiv e-prints*
- Lacasa F., Aghanim N., Kunz M., Frommert M., Apr. 2012, *MNRAS*, 421, 1982
- Lacasa F., Pénin A., Aghanim A., 2013, sub. to *MNRAS*
- Lagache G., Puget J.L., Dole H., Sep. 2005, *ARA&A*, 43, 727
- Lagache G., Bavouzet N., Fernandez-Conde N., et al., Aug. 2007, *ApJ*, 665, L89
- Laing R.A., Riley J.M., Longair M.S., Jul. 1983, *MNRAS*, 204, 151
- Lamarre J.M., Giard M., Pointecouteau E., et al., Nov. 1998, *ApJ*, 507, L5
- Leclercq F., Jasche J., Gil-Marín H., Wandelt B., May 2013, *ArXiv e-prints*
- Lewis A., Oct. 2011, *Journal of Cosmology and Astroparticle Physics*, 10, 26
- Lewis A., Bridle S., 2002, *Phys. Rev.*, D66, 103511
- Limousin M., Morandi A., Sereno M., et al., Aug. 2013, *Space Science Reviews*, 177, 155
- Lin Y.T., Partridge B., Pober J.C., et al., Apr. 2009, *ApJ*, 694, 992
- Lin Y.T., Shen Y., Strauss M.A., Richards G.T., Lunnan R., Nov. 2010, *ApJ*, 723, 1119
- Linde A.D., Oct. 1982, *Physics Letters B*, 116, 335
- Ma C.P., Fry J.N., 2002, In: Chen L.W., Ma C.P., Ng K.W., Pen U.L. (eds.) *AMiBA 2001: High-Z Clusters, Missing Baryons, and CMB Polarization*, vol. 257 of *Astronomical Society of the Pacific Conference Series*, 61
- MacTavish C.J., Ade P.A.R., Bock J.J., et al., Aug. 2006, *ApJ*, 647, 799
- Mak D.S.Y., Pierpaoli E., Dec. 2012, *Phys. Rev. D*, 86, 123520
- Mak D.S.Y., Pierpaoli E., May 2013, *Phys. Rev. D*, 87, 103518
- Maldacena J., May 2003, *Journal of High Energy Physics*, 2003, 013
- Manera M., Gaztañaga E., Jul. 2011, *MNRAS*, 415, 383
- Marscher A.P., 1996, In: Hardee P.E., Bridle A.H., Zensus J.A. (eds.) *Energy Transport in Radio Galaxies and Quasars*, vol. 100 of *Astronomical Society of the Pacific Conference Series*, 45

- McClure-Griffiths N.M., Pisano D.J., Calabretta M.R., et al., Apr. 2009, *Astrophysical Journal Supplement Series*, 181, 398
- Melin J.B., Bartlett J.G., Delabrouille J., Nov. 2006, *A&A*, 459, 341
- Miville-Deschênes M.A., Lagache G., Apr. 2005, *Astrophysical Journal Supplement Series*, 157, 302
- Mukhanov V.F., Chibisov G.V., May 1981, *Soviet Journal of Experimental and Theoretical Physics Letters*, 33, 532
- Navarro J.F., Frenk C.S., White S.D.M., Dec. 1997, *ApJ*, 490, 493
- Neugebauer G., Habing H.J., van Duinen R., et al., Mar. 1984, *ApJ*, 278, L1
- Neyrinck M.C., Sep. 2011a, *ArXiv e-prints*
- Neyrinck M.C., Dec. 2011b, *ApJ*, 742, 91
- Ostriker J.P., Vishniac E.T., Jul. 1986, *ApJ*, 306, L51
- Patanchon G., Ade P.A.R., Bock J.J., et al., Dec. 2009, *ApJ*, 707, 1750
- Peebles P.J.E., 1972, *Physical Cosmology*, Princeton Series in Physics, Princeton University Press
- Penco R., Mauro D., Sep. 2006, *European Journal of Physics*, 27, 1241
- Pénin A., Doré O., Lagache G., Béthermin M., Jan. 2012, *A&A*, 537, A137
- Pénin A., Lacasa F., Aghanim A., 2013, sub. to *MNRAS*
- Penzias A.A., Wilson R.W., Jul. 1965, *ApJ*, 142, 419
- Percival W.J., Reid B.A., Eisenstein D.J., et al., Feb. 2010, *Monthly Notices of the Royal Astronomical Society*, 401, 2148
- Peter P., Uzan J.P., 2005, *Cosmologie primordiale*, BELIN, First edn.
- Pietrobon D., Cabella P., Balbi A., de Gasperis G., Vittorio N., Jul. 2009, *MNRAS*, 396, 1682
- Pitrou C., Uzan J.P., Bernardeau F., Jul. 2010, *J. Cosmology Astropart. Phys.*, 7, 3
- Planck-Collaboration, Dec. 2011, *A&A*, 536, A18
- Planck collaboration, GASS collaborators, 2013, *A&A*, *to be submitted*
- Planck Collaboration, Abergel A., Ade P.A.R., et al., Dec. 2011a, *A&A*, 536, A24
- Planck Collaboration, Ade P.A.R., Aghanim N., et al., Dec. 2011b, *A&A*, 536, A7
- Planck Collaboration, Ade P.A.R., Aghanim N., et al., Dec. 2011c, *A&A*, 536, A19
- Planck Collaboration, Ade P.A.R., Aghanim N., et al., Dec. 2011d, *A&A*, 536, A1

Planck Collaboration VII, 2013, Submitted to A&A

Planck Collaboration XIX, 2013, Submitted to A&A

Planck Collaboration XV, 2013, Submitted to A&A

Planck Collaboration XVI, 2013, Submitted to A&A

Planck Collaboration XVII, 2013, Submitted to A&A

Planck Collaboration XVIII, 2013, Submitted to A&A

Planck Collaboration XX, 2013, Submitted to A&A

Planck Collaboration XXI, 2013, Submitted to A&A

Planck Collaboration XXIV, 2013, Submitted to A&A

Planck Collaboration XXIX, 2013, Submitted to A&A

Planck Collaboration XXVIII, 2013, Submitted to A&A

Planck Collaboration XXX., 2013, Submitted to A&A

Ponthieu N., Grain J., Lagache G., Nov. 2011, A&A, 535, A90

Press W.H., Schechter P., Feb. 1974, ApJ, 187, 425

Rasera Y., Alimi J.M., Courtin J., et al., Jun. 2010, In: Alimi J.M., Fuözfa A. (eds.)
American Institute of Physics Conference Series, vol. 1241 of American Institute of
Physics Conference Series, 1134–1139

Rees M.J., Sciama D.W., Feb. 1968, Nature, 217, 511

Reichardt C.L., Zahn O., Ade P.A.R., et al., Aug. 2009, ApJ, 701, 1958

Remazeilles M., Delabrouille J., Cardoso J.F., Feb. 2011, MNRAS, 410, 2481

Renaux-Petel S., Oct. 2009, Journal of Cosmology and Astroparticle Physics, 10, 12

Riess A.G., Macri L., Casertano S., et al., Jul. 2009, The Astrophysical Journal, 699,
539

Rindler W., 1956, MNRAS, 116, 662

Sachs R.K., Wolfe A.M., Jan. 1967, ApJ, 147, 73

Sefusatti E., Dec. 2005, Probing fundamental physics with large-scale structure: From
galaxy formation to inflation, Ph.D. thesis, New York University, New York, USA

Sefusatti E., Crocce M., Pueblas S., Scoccimarro R., Jul. 2006, Phys. Rev. D, 74, 023522

Sehgal N., Bode P., Das S., et al., Feb. 2010, ApJ, 709, 920

Serra P., Cooray A., May 2008, Phys. Rev. D, 77, 107305

- Sheth R.K., Tormen G., Sep. 1999, MNRAS, 308, 119
- Sievers J.L., Bond J.R., Cartwright J.K., et al., Jul. 2003, ApJ, 591, 599
- Silk J., Feb. 1968, ApJ, 151, 459
- Simpson F., Heavens A.F., Heymans C., Jun. 2013, ArXiv e-prints
- Smith K.M., Zaldarriaga M., Oct. 2011, MNRAS, 417, 2
- Smith R.E., Watts P.I.R., Sheth R.K., Jan. 2006, MNRAS, 365, 214
- Smoot G.F., Bennett C.L., Kogut A., et al., Sep. 1992, ApJ, 396, L1
- Spergel D., Goldberg D., Apr. 1999, Physical Review D, 59, 1, URL <http://link.aps.org/doi/10.1103/PhysRevD.59.103001>
- Starobinsky A.A., Nov. 1982, Physics Letters B, 117, 175
- Stoyanov J.M., 1987, Counterexamples in Probability, Probability and Statistics Series, Wiley & Sons, Second edn.
- Strukov I.A., Brukhanov A.A., Skulachev D.P., Sazhin M.V., Sep. 1992, MNRAS, 258, 37P
- Sunyaev R.A., Zeldovich Y.B., Mar. 1970, Comments on Astrophysics and Space Physics, 2, 66
- Sunyaev R.A., Zeldovich Y.B., Nov. 1972, Comments on Astrophysics and Space Physics, 4, 173
- Tegmark M., Blanton M.R., Strauss M.A., et al., May 2004, ApJ, 606, 702
- Thacker C., Cooray A., Smidt J., et al., May 2013, ApJ, 768, 58
- The Planck Collaboration, Apr. 2006, ArXiv Astrophysics e-prints
- Tinker J.L., Wetzel A.R., Aug. 2010, ApJ, 719, 88
- Toffolatti L., Argüeso Gomez F., de Zotti G., et al., Jun. 1998, MNRAS, 297, 117
- Toffolatti L., De Zotti G., Argüeso F., Burigana C., 1999, In: de Oliveira-Costa A., Tegmark M. (eds.) Microwave Foregrounds, vol. 181 of Astronomical Society of the Pacific Conference Series, 153
- Trindade A.M.M., Avelino P.P., Viana P.T.P., Apr. 2013, ArXiv e-prints
- Tucci M., Toffolatti L., De Zotti G., Martínez-González E., Aug. 2011, Astronomy & Astrophysics, 533, A57
- van den Bosch F.C., Yang X., Mo H.J., et al., Apr. 2007, MNRAS, 376, 841
- Viero M.P., Ade P.A.R., Bock J.J., et al., Dec. 2009, ApJ, 707, 1766

Weinberg D.H., Jan. 1992, MNRAS, 254, 315

Weinberg S., Jul. 1972, Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity, Wiley

Wick G.C., 1950, Physical Rev., 80, 268

Wilbanks T.M., Ade P.A.R., Fischer M.L., Holzapfel W.L., Lange A.E., Jun. 1994, ApJ, 427, L75

Wright E.L., Chen X., Odegard N., et al., Feb. 2009, Astrophysical Journal Supplement Series, 180, 283

Zheng Z., Berlind A.A., Weinberg D.H., et al., Nov. 2005, ApJ, 633, 791